Lehrstuhl für Bauinformatik und Geoinformationssysteme und Geodätisches Institut (gia)
Univ.-Prof. Dr.-Ing. Jörg Blankenbach

Lehr- und Forschungsgebiet Design Computation (DC)
Univ.-Prof. Dr. Jakob Beetz

Lehrstuhl für Energieeffizientes Bauen (E3D)
Univ.-Prof. Dr.-Ing. habil. Christoph van Treeck

Tagungsband des

36. Forum Bauinformatik

Proceedings of the 36th Forum Bauinformatik

24. – 26. September 2025

Herausgegeben von:
Baris Özcan, Hristo Vassilev,
Anne Göbels, Oliver Schulz,
Qirui Huang

© Peter Winandy

FORUMBAUINFORMATIK

Tagungsband des 36. Forum Bauinformatik 2025
Proceedings of the 36th Forum Bauinformatik 2025
Rheinisch-Westfälische Technische Hochschule Aachen
RWTH Aachen University

Herausgegeben von

Baris Özcan
Lehrstuhl für Bauinformatik und Geoinformationssysteme und Geodätisches Institut (gia)
RWTH Aachen University
Aachen, Germany

Hristo Vassilev
Lehrstuhl für Bauinformatik und Geoinformationssysteme und Geodätisches Institut (gia)
RWTH Aachen University
Aachen, Germany

Anne Göbels
Lehr- und Forschungsgebiet Design Computation (DC)
RWTH Aachen University
Aachen, Germany

Oliver Schulz
Lehr- und Forschungsgebiet Design Computation (DC)
RWTH Aachen University
Aachen, Germany

Qirui Huang
Lehrstuhl für Energieeffizientes Bauen (E3D)
RWTH Aachen University
Aachen, Germany

Layout & Satz
Qirui Huang

Verfügbar über das institutionelle Repositorium der RWTH Aachen University
DOI:10.18154/RWTH-2025-05890
Aachen 2025

Diese Veröffentlichung, ausgenommen Zitate und anderweitig gekennzeichnete Teile, ist unter der Creative-Commons-Lizenz CC BY 4.0 lizenziert.

Lizenzvertrag: Creative Commons Namensnennung CC BY 4.0
Das gedankliche Eigentum und die inhaltliche Verantwortung für die Beiträge liegt bei den jeweiligen AutorInnen.

Logo: Forum Bauinformatik | www.forum-bauinformatik.de
Umschlaggestaltung: Qirui Huang | RWTH Aachen University

Vorwort

Das Forum Bauinformatik dient seit 1989 regelmäßig dem Austausch im wissenschaftlichen Fachgebiet der Bauinformatik. Das 36. Forum findet 2025 zum zweiten Mal in Aachen statt und wird gemeinsam vom Lehrstuhl für Bauinformatik und Geoinformationssysteme und Geodätisches Institut (gia), Lehr- und Forschungsgebiet Design Computation (DC) und Lehrstuhl für Energieeffizientes Bauen (E3D) organisiert. Die Bauinformatik erfreut sich in den letzten drei Jahrzehnten einer stetig zunehmenden Aufmerksamkeit. In verschiedensten Teilbereichen treffen vielfältige methodische und technologische Entwicklungen auf zunehmend komplexere Herausforderungen.
Das Forum Bauinformatik bietet Wissenschaftlerinnen und Wissenschaftlern aller akademischen Grade aus dem DACH-Raum die Gelegenheit, Themen und Forschungsergebnisse zu diskutieren und abgeschlossene Projekte zu präsentieren. Da sich die Arbeit der Teilnehmerinnen und Teilnehmer des Forum Bauinformatik nicht auf den deutschsprachigen Raum beschränkt, sondern vielfältig auch mit Partnern sowie Kolleginnen und Kollegen aus der internationalen Forschungsgemeinschaft erfolgt, freuen wir uns besonders, sowohl deutsch- als auch englischsprachigen Beiträgen eine Plattform bieten zu können.
In diesem Konferenzband sammeln wir die insgesamt 59 Beiträge aus 12 Sessions zu den Themenbereichen:

- Building Information Modeling (BIM)
- Semantic Web & Linked Building Data
- Digital Twins & Smart Cities
- Visualization (AR, VR, XR)
- Structural Health Monitoring (SHM)
- User Behavior & Indoor Environmental Quality
- AI & Machine Learning
- Design Support
- Big Data & Data Mining
- Image Processing / Computer Vision
- Sustainability & Life Cycle Analysis (LCA)
- Sensors & Robotics
- Simulation & Optimization
- Geoinformation systems / Geoinformatics

Unser Dank gilt allen Autorinnen und Autoren sowie Betreuerinnen und Betreuern, die durch ihre Beiträge und Reviews zum Erfolg dieser Konferenz beigetragen haben. Wir wünschen allen Teilnehmenden eine spannende Konferenz mit inspirierenden Diskussionen und neuen Impulsen für ihre zukünftige Arbeit.

Baris Özcan, Hristo Vassilev, Anne Göbels, Oliver Schulz, Qirui Huang

Aachen, September 2025

Preface

Since 1989, the Forum Bauinformatik has regularly provided a platform for exchanging scientific research and ideas in the field of computing in civil engineering. The 36th Forum was held for the second time in Aachen in 2025 and was organized jointly by Chair of Computing in Civil Engineering and Geoinformation Systems and Geodetic Institute (gia), Chair of Design Computation (DC) and Institute of Energy Efficiency and Sustainable Building (E3D). Over the years, computing in civil engineering has gained increasing attention, with diverse methodological and technological developments addressing challenges of growing relevance and complexity.

Within the Forum Bauinformatik, scientists of all academic levels from the DACH region are given the opportunity to discuss topics and findings to present their work, engage in discussions, and showcase projects. As the work of the participants of the Forum Bauinformatik is not limited to the German-speaking region, but is also carried out with partners and colleagues from the international research community, we are particularly pleased to offer a platform for both German- and English-language contributions.

This conference volume compiles 59 contributions from 12 sessions, covering topics:

- Building Information Modeling (BIM)
- Semantic Web & Linked Building Data
- Digital Twins & Smart Cities
- Visualization (AR, VR, XR)
- Structural Health Monitoring (SHM)
- User Behavior & Indoor Environmental Quality
- AI & Machine Learning
- Design Support
- Big Data & Data Mining
- Image Processing / Computer Vision
- Sustainability & Life Cycle Analysis (LCA)
- Sensors & Robotics
- Simulation & Optimization
- Geoinformation systems / Geoinformatics

We extend our sincere thanks to all authors and supervisors for their contributions through papers and reviews. We wish all participants a productive conference filled with engaging discussions and insights for future work.

Baris Özcan, Hristo Vassilev, Anne Göbels, Oliver Schulz, Qirui Huang

Aachen, September 2025

Chronik des Forum Bauinformatik

1989	Hochschule für Architektur und Bauwesen Weimar
	Lehrstuhl für Informatik im Bauwesen
1990	Brandenburgische Technische Universität Cottbus
1991	Technische Universität Berlin
1992	Ruhr-Universität Bochum
	Lehrstuhl für Ingenieurinformatik im Bauwesen
1993	Technische Universität München
	Fachgebiet Bauinformatik
1994	Technische Hochschule Darmstadt
	Institut für Numerische Methoden und Informatik im Bauwesen
1995	Universität Hannover
	Institut für Bauinformatik
1996	Brandenburgische Technische Universität Cottbus
	Lehrstuhl für Bauinformatik
1997	Technische Universität Dresden
	Lehrstuhl für Computeranwendungen im Bauwesen
1998	Bauhaus-Universität Weimar
	Lehrstuhl für Informatik im Bauwesen
1999	Technische Hochschule Darmstadt
	Institut für Numerische Methoden und Informatik im Bauwesen
2000	Technische Universität Berlin
	Theoretische Methoden der Bau- und Verkehrstechnik, Institut für Bauingenieurwesen
2001	Technische Universität München
	Lehrstuhl für Bauinformatik
2002	Ruhr-Universität Bochum
	Lehrstuhl für Ingenieurinformatik im Bauwesen
2003	Universität Hannover
	Institut für Bauinformatik
2004	Technische Universität Braunschweig
	Institut für Computeranwendungen im Bauingenieurwesen
2005	Brandenburgische Technische Universität Cottbus
	Lehrstuhl für Bauinformatik
2006	Bauhaus-Universität Weimar
	Professur Informatik im Bauwesen
2007	Technische Universität Graz
	Institut für Bauinformatik

2008	Technische Universität Graz
	Institut für Informatik
2009	Universität Karlsruhe
	Institut für Industrielle Bauproduktion und Entwerfen
2010	Technische Universität Berlin
	Fachgebiet Bauinformatik, Institut für Bauingenieurwesen
2011	University College Cork
	Informatics Research Unit for Sustainable Engineering
2012	Ruhr-Universität Bochum
	Lehrstuhl für Informatik im Bauwesen
2013	Technische Universität München
	Lehrstuhl für Computergestützte Modellierung und Simulation
	Lehrstuhl für Architekturinformatik
2014	Technische Universität Darmstadt
	Institut für Numerische Methoden und Informatik im Bauwesen
2015	Rheinisch-Westfälische Technische Hochschule Aachen
	Geodätisches Institut, Lehrstuhl für Bauinformatik & Geoinformationssysteme
2016	Leibniz Universität Hannover
	Institut für Risiko und Zuverlässigkeit
2017	Technische Universität Dresden
	Institut für Bauinformatik
2018	Bauhaus-Universität Weimar
	Professur Informatik im Bauwesen
2019	Technische Universität Berlin
	Fachgebiet Bauinformatik und Fachgebiet Systemtechnik baulicher Anlagen
2021	Technische Hochschule Darmstadt
	Institut für Numerische Methoden und Informatik im Bauwesen
2022	Technische Universität München
	Lehrstuhl für Computergestützte Modellierung und Simulation
	Lehrstuhl für Architekturinformatik
2023	Ruhr-Universität Bochum
	Lehrstuhl für Ingenieurinformatik im Bauwesen
2024	Technische Universität Hamburg
	Institut für Digitales und Autonomes Bauen
2025	Rheinisch-Westfälische Technische Hochschule Aachen
	Lehrstuhl für Bauinformatik und Geoinformationssysteme und Geodätisches Institut
	Lehr- und Forschungsgebiet Design Computation
	Lehrstuhl für Energieeffizientes Bauen

Inhaltsverzeichnis

Kopplung von Building Information Modeling mit Behavior Trees als Grundlage für automatisierte Bauwerksinspektionen **1**
Aditya Tandon, Jan Stührenberg and Saurabh Sawant

Von Materialtests zum automatisierten Monitoring von 3D-gedruckten Lehmbauwerken **9**
Jasper Vollmert, P. Peralta and A. Alatassi

Mobile Application to Support Facade Renovation **17**
Oliver Quintel, Hendrik Martin Würz, Michel Krämer, Eva Klien and Arjan Kuijper

Structured Databases for Heritage BIM: Enhancing Digital Documentation and Restoration Workflows **25**
Laura Marcela Valderrama Niño and Sabrina Becker

BIM and Semantic Web: Key Enablers for Nuclear Knowledge Management? **33**
Pascal Mosler, Ahmed Han Karapinar, Benedikt Malte Hauke Broich and Theo Kastner-Guhl

Verbesserung des Informationsflusses in der Infrastrukturplanung: Ein systemtheoretischer Ansatz **41**
Bastian Wons and Martin Wogan

Building Knowledge Representation for Computer Scientists: Insights from Teaching . **49**
Diellza Elshani, Daniel Hernandez, Steffen Staab and Thomas Wortmann

Integrated Real-Time Monitoring of Sewer Systems via SPC and Graph-Based Validation **57**
Jun Li and Finia Buitkamp

Towards Semantic Interoperability Between IFC and CityGML Using Semantic Web Technologies **66**
G.R. Sirigireddy and J. Krischler

BIM- and GIS-data linkage proposals using Explainable AI (XAI) framework **74**
G. Nikolaev and J. Krischler

Discrete Event Simulation (DES)-Based Abstraction of Construction State Machines and Robotic Simulations for Site-Wide Optimization Using BIM and Gazebo **82**

Ankita Maurya

Unsupervised Semantic Segmentation of Cracks in Concrete Specimens **90**
Amir Sadrnia, Mohamed S. H. Alabassy and Andrea Osburg

Hygrothermal Performance of Clinker Bricks with Perforations: A Comparative Simulation Study **98**
Andrea Thiele, Daniel Kastner, Oliver Abele and Jérôme Frisch

Overcoming Stakeholder Knowledge Gaps: A HITL-Oriented Ontology Mapping Framework for the Construction Industry **107**
Jakob Martin, Timur Weilbach-Eyüboglu and Vedant Girish Dalvi

Targeted Data Exchange for Efficient MEP Openings in BIM using IDS **115**
Timur Weilbach-Eyüboglu, Jakob Martin and Vedant Girish Dalvi

Qualitative Analyse räumlicher Bedarfe und Anforderungen infolge von Homeoffice . . **123**
Anna Langenbeck, Marc Syndicus, Christoph van Treeck and Jérôme Frisch

Utilization of deep autoencoders for improving semantic segmentation of 3D point clouds **131**
Marcel Aubart and Hristo Vassilev

Untersuchung zur Akzeptanz und Nutzung von Energie- und Komfortmonitoringsystemen in privaten Haushalten **139**
Lisa von Wittenhorst zu Sonsfeld and Elisabeth Beusker

The Role of AI Agents in Construction Project Management **147**
Anja Brelih, Marko Robnik Šikonja and Robert Klinc

Automated Extraction of Information Requirements from Building Regulations Using LLMs and Ontologies **155**
Sven Zentgraf and Leonie Zimmermann

An ontology-based approach for sewer infrastructure management **163**
Raquel Valles Gómez, Luis Jamora, Peter Gölzhäuser, Sabine Hartmann, Amirhossein Babaei Ravandi, Hendrik Morgenstern and Katharina Klemt-Albert

Digital Heritage in VR: 360° Media, Theological Insights and Interactive Artifact Exploration – The Case of Tel Lachish **171**
Djamel Berkaoui, Christina Kumpmann, Koen Castermans, Raymond Leonardo Chandra, Patrick Querl and Heribert Nacken

Datenanalysekonzept für die Anwendung von BIM in der Bauausführung von Infrastrukturprojekten **179**

Tim Noack, Philip Markert, Björn Buuck, Caroline Barr, Al-Hakam Hamdan, Chongjie Kang and Steffen Marx

Virtual Masonry Shear Wall Experiments: Enhancing Understanding of Nonlinear Behavior through Virtual Simulation **187**
Raymond Leonardo Chandra, Stephan Dreyer, Florian Kolisch, Koen Castermans, Djamel Berkaoui, Patrick Querl, Heribert Nacken and Sven Klinkel

Virtual Structural Analysis of a Shaking Table for Earthquake Simulations **195**
Raymond Leonardo Chandra, Stephan Dreyer, Florian Kolisch, Koen Castermans, Djamel Berkaoui, Patrick Querl, Heribert Nacken and Sven Klinkel

Parametrische Simulation von Linienbaustellen für KI-gestützte Ressourcenplanung . **203**
Anna-Maria Mehringer, Jonas Wiederer, Maik Görgner and Simon Höng

Graphbasierte Visualisierung von Produktdaten der Technischen Gebäudeausrüstung **211**
Erich Dierkes

Construction site obstacle detector assisted by large language models **219**
Miguel Guzmán-Merino and Dennis Grauer

Short-term HVAC load prediction using Temporal Graph Convolutional Networks **227**
Dennis Grauer, Miguel Guzman-Merino and Nils Soeren Krause

Prototypische Implementierung einer Konstruktions- und Regelsprache für Geometrie **235**
Jan Hildebrand and Marian Kurzawa

Building a European-Level Energy Registry **243**
Mahsa Faraji Shoyari, Jessy Matar and Markus Duchon

Transforming IFC Models into a Structured 3D Format for Multimodal Localization from 2D Images in a Web-Based Tool **251**
Robin Ov, Ayman Soultana, Angelina Aziz and Roberto Perez Martinez

Reconstruction of IFC Models Using Voxelization for Spatial Analysis **259**
Cemre Karaoglan

An Evaluation of Zero-shot Classification of Multi-Modal Foundation Models for Point Clouds **267**
Nils Sören Krause, Miguel Guzman-Merino and Dennis Grauer

Evaluating the DIN EN ISO 11855 1D-RC-approach for modeling temperature profiles of pipe embedding façade systems **276**
Oliver Abele, Hannah Riefer, Jérôme Frisch and Christoph van Treeck

Benchmarking Large Language Models of BIM: Key Requirements for Domain-Specific Evaluation Datasets **284**
Abdelrahman Abouhadid and Clemens Kujat

Improving Data Robustness in Environmental Monitoring: Lessons from a Green Façade Study **293**
Harald Lesan

Enhancing Trust in Machine Learning-Based Architectural Design Space Exploration Through Uncertainty Visualization, a Pilot Study **301**
Max Zorn and Thomas Wortmann

Comprehensive Georeferencing Support for all Official Versions of the IFC Standard . **309**
Štefan Jaud, Paul-Christian Max, Torben Pullmann and Andreas Geiger

Integriertes und skalierbares Planungssystem auf Basis von GIS- und Bodeninformationen in der Infrastrukturplanung **317**
Maik Görgner, Simon Hoeng, Anna-Maria Mehringer and Jonas Wiederer

TopoBEM: Using Non-Manifold Topology and Building Energy Simulation to Facilitate Evidence- Based Decision Making in Architectural Design **325**
Malte Junge

Ontology-Driven GPT-4 Question Answering on Construction Standards **333**
Merna Emara

Persistence and linking of geometric-topological data models in BIM and GIS **341**
Sophie Lipskaya, Robert Jürgensen and Maximilian Sternal

Image Captioning for Building Damage Using Deep Learning: Explaining Context in Overlapping Structural Defects **349**
Vedant Girish Dalvi, Jakob Martin and Timur Weilbach-Eyüboglu

BIKE-SCAN: A GIS Extension for Bicycle Infrastructure Safety Analysis Based on Street-Level Images **358**
Michael Disser

End-to-End Parametric Workflow for Bridge Design with Automated Cost Analysis and Generative Modeling **366**
Jonas Wiederer, Simon Hoeng, Maik Görgner and Anna-Maria Mehringer

Automating GIS-Derived Site Plans for Digital Building-Permit **374**
Fatma Ahmad and Ralf Becker

Prozessoptimierung in der BIM-basierten Tragwerksplanung: Entwicklung eines IFC-SAF-Konverters **382**
Niklas Maczyewski, Dominik Schlütter, Jérôme Frisch and Christoph van Treeck

An API-Driven System for Real-Time Safety Monitoring on Construction Sites: Integrating Machine Learning and Industry Foundation Classes **390**
Mukta Nisal and Xiangyu Wang

A Digital Twin Prototype for a Model-Scale Robotic Vehicle: Architecture, Implementation, and Integration Perspectives **398**
Fritz Trede, Mohammad Alallo, Youssef Aknin and Vikram Hawladar

Ontology-Augmented Multi-Agent Reinforcement Learning for Enhancing EV Charging Network Recovery **406**
Mohamed Abdelfattah

Knowledge-Enhanced Cobot Reasoning Assistant for Robot Manipulation **414**
Ruibo Tang, Clara-Larissa Lorenz, Jérôme Frisch and Christoph van Treeck

A system for querying spatial relations between technical drawings and geometric data using SPARQL **422**
Fabian Rolke

Datengetriebener Digital Twin für das Facility Management: Integration von Building Information Modeling, Multi-Modell-Container und Large Language Models **430**
Thujiethaa Thuraisingam, Nijanthan Mohan and Rolf Groß

Matrix-Based Approach to Optimizing Disassembly Sequence Planning for Aluminum-Framed Windows **438**
Begüm Aktas

Regelbasierte Modellprüfung im Krankenhausbau: Entwicklung eines Regelsatzes zur automatisierten Qualitätskontrolle **446**
Sabine Hartmann, Linus Kloesgen, Hendrik Morgenstern and Katharina Klemt-Albert

Synthetic Image Data from 3D Software Environments and Game Engines for Training Machine Learning Models in Damage Detection **454**
Mahmoud Moubarak and Baris Özcan

Automatisierte Fehlererkennung im Schmelzschichtverfahren mittels künstlicher Intelligenz **461**
Thamer Al-Zuriqat, Mahmoud Noufal, Adel Alatassi and Jasper Vollmert

Concept for digitally AI-readable buildings for urban mining based on digital building logbooks using large language models **470**
Heiß Marcel, Benedikt Kandler and Uwe Rüppel

Kopplung von Building Information Modeling mit Behavior Trees als Grundlage für automatisierte Bauwerksinspektionen

Aditya Tandon, Jan Stührenberg and Saurabh Sawant
Institut für Digitales und Autonomes Bauen, Technische Universität Hamburg, Hamburg, Deutschland
E-mail(s): aditya.tandon@tuhh.de

Abstract: Das Monitoring und regelmäßige Wartungsinspektionen von Gebäuden und Infrastrukturbauwerken sind unerlässlich, um Qualitätsstandards aufrechtzuerhalten, die die strukturelle Sicherheit und Integrität gewährleisten. Da manuelle Inspektionen zeitaufwändig und teuer sein können, wird der Einsatz von Robotern erforscht, um Gebäudeinspektionen zu automatisieren und Wartungsprozesse zu optimieren. Die Einführung robotergestützter Inspektionsmethoden wird jedoch durch eine fehlende Kopplung mit der Bauwerksinformationsmodellierung (Building Information Modeling, BIM) für eine situationsbezogene Navigation sowie eine mangelnde Anpassungsfähigkeit der Inspektionsplanung erschwert. In diesem Beitrag wird ein BIM-basiertes Framework für automatisierte Bauwerksinspektionen vorgestellt, um die genannten Herausforderungen zu bewältigen. Räumliche und semantische Daten werden aus BIM-Modellen extrahiert, um die Roboternavigation zu verbessern und zu inspizierende Gebäudeelemente zu identifizieren. Behavior Trees werden für eine anpassungsfähige und automatisierte Planung und Ausführung von Roboterinspektionsaufgaben genutzt. Das BIM-basierte Framework für automatisierte Bauwerksinspektionen wird auf vierbeinigen Robotern implementiert und durch die Durchführung von Inspektionsaufgaben in Bürogebäuden validiert. Die Ergebnisse zeigen, dass die Kopplung von BIM und Behavior Trees die Genauigkeit und Effizienz von automatisierten Gebäudeinspektionen verbessert. Zusammenfassend lässt sich sagen, dass das BIM-basierte Framework für automatisierte Bauwerksinspektionen einen Schritt in Richtung der Einführung von robotergestützter Gebäudeinspektion in der Praxis darstellt.

Keywords: Gebäudeinspektionen, Bauwerksinspektionen, Bauwerksinformationsmodellierung (BIM), Behavior Trees, vierbeinige Roboter, autonome Navigation

DOI: 10.18154/RWTH-CONV-254872. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Einführung

Regelmäßige Inspektionen von Gebäuden sind für die Aufrechterhaltung von Qualitätsstandards, die die strukturelle Sicherheit und Integrität gewährleisten, unerlässlich [1]. Gebäudeinspektionen werden üblicherweise manuell durchgeführt, wobei geschulte Inspektoren visuelle und strukturelle Bewertungen vornehmen [2]. Manuelle Inspektionen sind jedoch oft zeitaufwändig, arbeitsintensiv und

gefährlich und neigen zu subjektiven Qualitätsbewertungen [3]. Deshalb haben sich mobile Roboter als praktikable Lösung für die Automatisierung von Wartungsinspektionen von Gebäuden erwiesen. Sie bieten das Potenzial für sicherere, häufigere und objektivere Qualitätsbewertungen im Vergleich zu herkömmlichen manuellen Inspektionen [4]. Die Bauwerksinformationsmodellierung (Building Information Modeling, BIM) wird zunehmend eingesetzt, um standardisierte digitale Gebäudemodelle über den gesamten Lebenszyklus von Gebäuden zu nutzen [5]. BIM-Modelle enthalten sowohl semantische als auch geometrische Informationen, die für die Lokalisierung von Robotern innerhalb von Gebäuden, die Identifizierung zu inspizierender Gebäudeelemente, die Planung optimaler Pfade zu den Gebäudeelementen und die Durchführung von Inspektionsaufgaben genutzt werden können [6].

Kim und Peavy [7] haben ein Framework entwickelt, das BIM-Daten in das Universal Robot Description Format konvertiert, um Inspektionen zu planen und durchzuführen. Die Validierung wird jedoch ausschließlich in simulierten Umgebungen durchgeführt. Chen et al. [8] haben ein BIM-basiertes System für automatisierte Gebäudeinspektionen mit Robotern unter Annahme einer geschlossenen Welt vorgeschlagen. Insgesamt wird die Interoperabilität zwischen BIM und Robotern nur selten berücksichtigt, und die Anpassungsfähigkeit an reale dynamische Umgebungen oder die Neuplanung von Inspektionen ist begrenzt [9]. Darüber hinaus können Abweichungen zwischen dem geplanten Zustand, der in BIM-Modellen widergespiegelt wird, und dem tatsächlichen Zustand die Aufgaben- und Bewegungsplanung von Robotern bei der Automatisierung von Gebäudeinspektionen beeinträchtigen. Um die Anpassungsfähigkeit von Robotern in dynamischen Umgebungen zu verbessern, haben Behavior Trees an Popularität gewonnen. Der hierarchische Aufbau von Behavior Trees ermöglicht eine adaptive und reaktive Aufgabenplanung in Verbindung mit flexiblen Strategien zur Aufgabenausführung in unsicheren und dynamischen Umgebungen [10].

In diesem Beitrag wird ein BIM-basiertes Inspektionsframework präsentiert, das BIM und Behavior Tree (BT)-Kontrollstrukturen verbindet, um automatisierte Gebäudeinspektionen mit vierbeinigen Robotern durchzuführen. Zu inspizierende Gebäudeelemente werden aus BIM-Modellen extrahiert. Um zu den Gebäudeelementen zu navigieren, werden künstliche Marker eingesetzt, um die BIM-Modelle an Karten für die Roboternavigation auszurichten. Inspektionsaufträge werden als BTs generiert, basierend auf der Art der Gebäudeelemente und den Fähigkeiten der Roboter zur Durchführung von Inspektionen.

Der Rest des Beitrags ist wie folgt gegliedert. Zunächst wird der Entwurf des BIM-basierten Inspektionsframeworks vorgestellt. Daraufhin werden die Implementierung und Validierung des Frameworks beschrieben und die Ergebnisse der Validierungstests diskutiert. Abschließend werden die Ergebnisse zusammengefasst und mögliche zukünftige Forschungsrichtungen vorgeschlagen.

2 Ein BIM-basiertes Inspektionsframework mit BT-Kontrollstrukturen

In diesem Abschnitt wird das Design des BIM-basierten Inspektionsframeworks vorgestellt. Zunächst wird die Extraktion von Gebäudeelementen aus BIM-Modellen für die Generierung von Inspektionsaufgaben vorgestellt, gefolgt von der Ausrichtung von BIM-Modellen mit Karten für die Roboternavigation und der automatischen Generierung von Behavior-Trees aus Gebäudeelementen.

2.1 Extraktion von Gebäudeelementen aus BIM-Modellen

Als Grundlage für zuverlässige und präzise Inspektionen in Gebäuden werden Informationen, wie etwa Abmessungen und Positionen, über spezifische zu inspizierende Gebäudeelemente („Zielelemente“) aus BIM-Modellen extrahiert. Inspektionsposen werden in festgelegtem Abstand zu den Zielelementen definiert und so ausgerichtet, dass sie die Zielelemente frontal erfassen und somit eine optimale Inspektion ermöglichen. Die Inspektionspose ${}^{\mathrm{B}}_{\mathrm{I}}\mathbf{P}$ im Bezugssystem des BIM-Modells **B**, für das n-te Gebäudeelement vom Typ **T**, wird folgendermaßen berechnet:

$$ {}^{\mathrm{B}}_{\mathrm{I}}\mathbf{P}_{\mathrm{T}_n} = \mathbf{H}_{\mathrm{T}}\, {}^{\mathrm{B}}_{\mathrm{E}}\mathbf{P}_{\mathrm{T}_n} \in SE(3) \tag{1} $$

Dabei bezeichnet ${}^{\mathrm{B}}_{\mathrm{I}}\mathbf{P}_{\mathrm{T}_n}$ die Inspektionspose für das n-te Element vom Typ **T** in **B**, $\mathbf{H}_{\mathrm{T}}$ die homogene Transformationsmatrix, die den festen Abstand und die Ausrichtung für Elemente des Typs T enthält, und ${}^{\mathrm{B}}_{\mathrm{E}}\mathbf{P}_{\mathrm{T}_n}$ die Pose des Gebäudeelements für das n-te Element vom Typ T in **B**.

2.2 Ausrichtung von BIM-Modellen mit Karten für die Roboternavigation

Die Zielelemente sind im Bezugssystem des BIM-Modells **B** definiert. Bei herkömmlichen Roboternavigationsansätzen werden jedoch Karten zur Lokalisierung und Pfadplanung eingesetzt, die üblicherweise in einem von **B** separaten Koordinatensystem **M** definiert sind. Daher ist eine Transformation ${}^{\mathrm{B}}\mathbf{T}_{\mathrm{M}}$ erforderlich, um das Bezugssystem des BIM-Modells **B** und das Kartenkoordinatensystem **M** auszurichten.

Das BIM-basierte Inspektionsframework berechnet die starre Transformation ${}^{\mathrm{B}}\mathbf{T}_{\mathrm{M}}$ anhand mehrerer Korrespondenzen von leicht zu erkennenden Referenzmarken. Als Referenzmarken werden AprilTags verwendet [11], die an denselben Stellen im realen Gebäude und im entsprechenden BIM-Modell platziert werden. Durch die Platzierung "virtueller" AprilTags im BIM-Modell können die Positionen der AprilTags in Bezug auf **B** abgerufen werden. Die Beobachtung der in realen Gebäuden platzierten AprilTags mit an Robotern montierten Kameras ermöglicht die Berechnung von Transformationen von den Robotern zu den AprilTags. Kombiniert man die Transformationen mit den aktuellen Posen der Roboter in der Karte, erhält man die Posen der AprilTags in Bezug auf **M**. Da die AprilTags ID-Nummern enthalten, lassen sich die entsprechenden AprilTags in BIM-Modellen und in realen Gebäuden schnell und eindeutig identifizieren. Während die Roboter Gebäude durchqueren und AprilTags beobachten, werden Korrespondenzen von realen und virtuellen AprilTags erfasst. Wenn das erste AprilTag beobachtet wird, wird die erste Korrespondenz erfasst, und die starre Transformation ${}^{\mathrm{B}}\mathbf{T}_{\mathrm{M}}$ kann wie folgt berechnet werden:

$$ {}^{\mathrm{B}}\mathbf{T}_{\mathrm{M}} = {}^{\mathrm{B}}\mathbf{T}_{\mathrm{A}_i}\, {}^{\mathrm{M}}_{t}\mathbf{T}^{-1}_{\mathrm{A}_i} \in SE(3) \quad \text{mit} \quad {}^{\mathrm{M}}_{t}\mathbf{T}_{\mathrm{A}_i} = {}^{\mathrm{M}}_{t}\mathbf{T}_{\mathrm{R}}\, {}^{\mathrm{R}}_{t}\mathbf{T}_{\mathrm{A}_i} \tag{2} $$

Dabei bezeichnet ${}^{\mathrm{B}}\mathbf{T}_{\mathrm{A}_i}$ die Pose des AprilTags mit der ID *i* im Bezugsystem des BIM-Modells **B**, ${}^{\mathrm{M}}_{t}\mathbf{T}^{-1}_{\mathrm{A}_i}$ die Transformationsschätzung des AprilTags mit der ID *i* in das Kartenkoordinatensystem **M** zum Zeitpunkt *t*, ${}^{\mathrm{M}}_{t}\mathbf{T}_{\mathrm{R}}$ die Pose des Roboters in **M** zum Zeitpunkt *t* und ${}^{\mathrm{R}}_{t}\mathbf{T}_{\mathrm{A}_i}$ die Transformation vom durch die Kamera beobachteten AprilTag mit der ID *i* in das Roboterkoordinatensystem **R** zum Zeitpunkt *t*. Die Genauigkeit von ${}^{\mathrm{M}}_{t}\mathbf{T}_{\mathrm{A}_i}$ und folglich auch die Genauigkeit von ${}^{\mathrm{B}}\mathbf{T}_{\mathrm{M}}$ wird jedoch durch Messrauschen und die Bewegung des Roboters beeinflusst und können zu einer ungenauen

Ausrichtung des BIM-Modells und der Karte führen. Um eine genaue Transformation ${}^{\mathrm{B}}\mathbf{T}_{\mathrm{M}}$ zu erhalten, werden *m* Detektionen von AprilTags in aufeinanderfolgenden Kamerabildern berücksichtigt, um die mittlere starre Transformation ${}^{\mathrm{M}}\overline{\mathbf{T}}_{\mathrm{A}_i}$ zu berechnen. Die Varianz dieser Transformationen wird auf Basis des Root Mean Square Error (RMSE) bewertet. Transformationen geringer Qualität, die möglicherweise zu Fehlausrichtungen führen, werden identifiziert, wenn der RMSE den Schwellwert τ_{RMSE} überschreitet, und von weiteren Berechnungen ausgeschlossen.

Zusätzlich wird die mittlere Transformation ${}^{\mathrm{B}}\overline{\mathbf{T}}_{\mathrm{M}}$ anhand mehrerer AprilTag-Korrespondenzen berechnet, um mögliche Mess- und Platzierungsfehler der AprilTags auszugleichen und eine genaue Ausrichtung zu gewährleisten. Die Transformation ${}^{\mathrm{B}}\overline{\mathbf{T}}_{\mathrm{M}}$ wird neu berechnet, sobald eine neue AprilTag-Korrespondenz mit einer hochwertigen Transformation ${}^{\mathrm{M}}\overline{\mathbf{T}}_{\mathrm{A}_i}$ akzeptiert wurde.

2.3 Automatisierte Generierung von BTs aus Gebäudeelementen

Eine zuverlässige und genaue Planung von Inspektionsaufgaben erfordert Strukturen, die den Ausführungsablauf von Robotern dynamisch adaptieren. In dieser Studie werden Behavior Trees aufgrund ihrer Modularität, Anpassungsfähigkeit und Wiederverwendbarkeit verwendet. Im Wesentlichen werden die Zielelemente aus Abschnitt 2.1 analysiert, um die Art und den Ort der für die Gebäudeinspektion erforderlichen Inspektionsaufgaben zu bestimmen. Für jedes Zielelement wird ein eindeutiger Satz von Inspektionsaufgaben definiert, der auf funktionalen und betrieblichen Anforderungen basiert. Die Inspektionsaufgaben werden den spezifischen Roboterfähigkeiten zugeordnet, um sicherzustellen, dass die Roboter die Anforderungen erfüllen können. Auf der Grundlage der Zuordnung von Inspektionsaufgaben und Roboterfähigkeiten werden BT-Subtrees generiert, die die Logik, Parameter und Ausführungssequenzen enthalten.

In dieser Studie werden der Navigations-Subtree und der visuelle Inspektions-Subtree entwickelt. Der in Abbildung 1 dargestellte Navigations-Subtree erleichtert die autonome Planung und Ausführung von Navigationspfaden und die Fehlerbehebung im Falle von Navigationsfehlern. Der Navigations-Subtree wertet die Bedingungen "Zielpose erhalten" und "Zielpose erreicht" aus, bevor die Navigation erfolgt. Ziele werden mithilfe von "Ziel planen" geplant und "Zum Ziel navigieren" ausgeführt.

Der Subtree zur visuellen Inspektion, dargestellt in Abbildung 2 führt eine Reihe von visuellen Inspektionen innerhalb der Betriebsumgebung aus. Nach der Auswertung der Bedingung zum Start der Inspektion aktiviert der Subtree die für die visuelle Inspektion notwendigen Module (z.B. Kamera und Objekterkennung) und erfasst Bilder zur Dokumentation.

Um größere Abfolgen von Inspektionsaufgaben zu erstellen, werden der Navigations-Subtree und der Subtree zur visuellen Inspektion in Steuerstrukturen für die Ausführung von BTs miteinander verknüpft. Dadurch können Roboter zu Inspektionspunkten navigieren, visuelle Inspektionen durchführen und zu nachfolgenden Inspektionspunkten weiterziehen. Im folgenden Abschnitt werden die Implementierung und Validierung des BIM-basierten Inspektionsframeworks beschrieben und die Ergebnisse diskutiert.

3 Implementierung und Validierung des Inspektionsframeworks

In diesem Abschnitt wird die Implementierung und Validierung des BIM-basierten Inspektionsframeworks vorgestellt. Schließlich werden die Ergebnisse der Validierungstests vorgestellt und diskutiert.

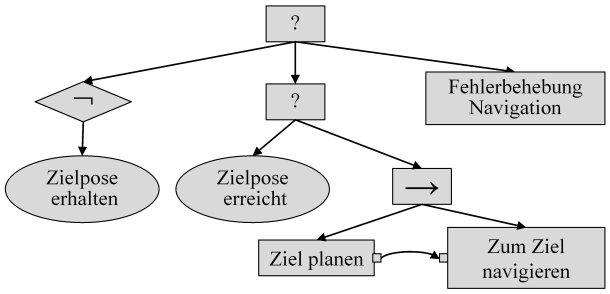


Abbildung 1: Navigations-Subtree

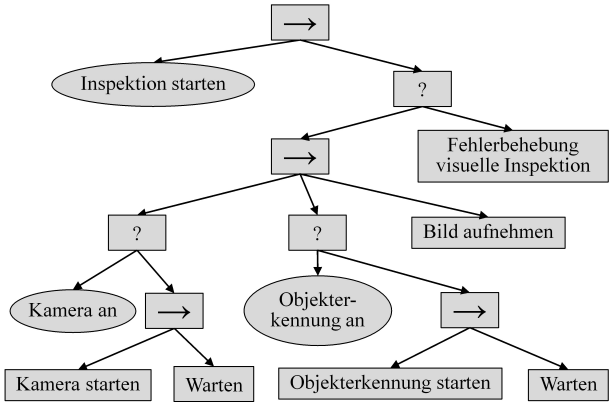


Abbildung 2: Visueller Inspektions-Subtree

3.1 Implementierung

Um die Extraktion der Gebäudeelemente aus BIM-Modellen zu ermöglichen, wird das offene und standardisierte Industry Foundation Classes (IFC)-Schema verwendet. Zur Implementierung des BIM-basierten Inspektionsframeworks wird die Open-Source-Bibliothek IfcOpenShell verwendet, die das Parsen von IFC-Dateien ermöglicht. Die Extraktion von Gebäudeelementen erfolgt über die Python-API der IfcOpenShell-Bibliothek. Sobald die Zielelemente ermittelt sind, werden entsprechende Inspektionsposen für die Roboter in vordefinierten Abständen zu den Zielelementen in **B** berechnet.

Um das BIM-Modell und die Karten zur Navigation auszurichten, wird eine benutzerdefinierte, mit Autodesk Revit kompatible AprilTag-Familie verwendet, die im BIM-Modell platziert wird. Aus dem BIM-Modell werden die Tags als *IfcBuildingElementProxy*-Elemente in das IFC-Format exportiert, und die Posen der AprilTags ${}^{\mathrm{B}}\mathbf{T}_{\mathrm{A}_i}$ werden mithilfe der IfcOpenShell-Bibliothek automatisch extrahiert. Im realen Gebäude werden die AprilTags auf Papier gedruckt und entsprechend den Positionen im BIM-Modell platziert, und die Bibliothek *apriltag_ros* wird verwendet, um die Tags im realen Gebäude zu erkennen. Das Optimierungsproblem der Berechnung der mittleren Transformationen ${}^{\mathrm{M}}\overline{\mathbf{T}}_{\mathrm{A}_i}$ und ${}^{\mathrm{B}}\overline{\mathbf{T}}_{\mathrm{M}}$ ist in Python unter Verwendung der *scipy.optimize.least_squares* Bibliothek implementiert.

Die Kontrollstrukturen für die Ausführung von BTs werden mithilfe von Python im Robot Operating System (ROS) implementiert. Zur Generierung von BTs für Inspektionsmissionen wird das ros_bt_py-Framework [12] erweitert, um benutzerdefinierte Knoten zu erstellen. Die Navigations- und visuellen Inspektions-Subtrees werden über eine Programmierschnittstelle (API) mit einer vordefinierten logischen Struktur implementiert. Die Subtrees werden mithilfe von ROS-Serviceaufrufen miteinander verkettet. Für weitere Informationen zur Roboternavigation sei auf [3] verwiesen.

3.2 Validierung

Die Validierungstests werden in einer Büroumgebung mit den Abmessungen 39 m x 16 m und einem entsprechenden BIM-Modell durchgeführt (Abbildung 3). Die Tests dienen dazu, die Genauigkeit und Zuverlässigkeit des BIM-basierten Inspektionsframeworks zu validieren. Die Genauigkeit wird durch Messung der Abstände zwischen den geplanten sowie den von den Robotern erreichten Inspektionsposen und Berechnung des RMSE bewertet. Die Zuverlässigkeit des BIM-basierten Inspektionsframeworks wird anhand der Fähigkeit bewertet, Inspektionsaufgaben automatisch ohne

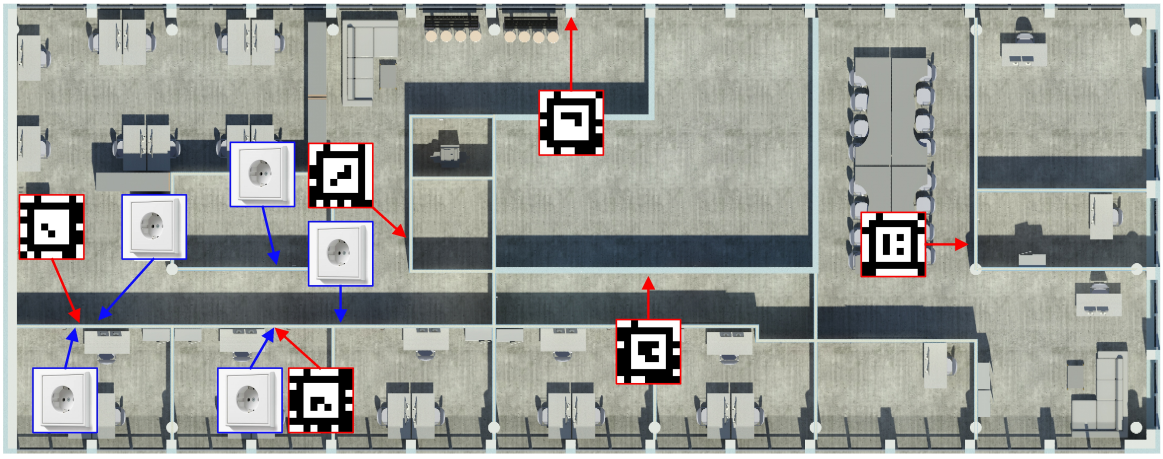


Abbildung 3: BIM-Modell der Büroumgebung mit Steckdosen und AprilTags

menschliches Eingreifen durchzuführen, d. h. (i) Navigation zu den Inspektionsposen, (ii) Ausführung der BT-Strukturen und (iii) erfolgreiche Erkennung und Dokumentation von Zielelementen.

Die Validierungstests werden wie folgt durchgeführt: Elektrische Steckdosen werden als Zielelemente für visuelle Inspektionen ausgewählt. Das in das IFC-Format exportierte BIM-Modell der Büroumgebung wird nach dem entsprechenden IFC-Element IfcOutlet durchsucht. Dann werden die Inspektionsmissionen, bestehend aus Navigations- und Sichtprüfungsaufgaben, geplant. Die Inspektionsposen werden mit 1 m Abstand zu den Steckdosen geplant. Für eine genaue Ausrichtung des BIM-Modells und der Karten werden sechs AprilTags strategisch an den Wänden der Büroumgebung und im BIM-Modell platziert (Abbildung 3).

Die Validierungstests werden mit einem vierbeinigen Roboter, dem in Abbildung 4 gezeigten "Intelligent Documentation Gadget" (IDOG), durchgeführt, der bereits in früheren Studien erfolgreich zur Durchführung von Inspektionen zur strukturellen Gesundheitsüberwachung eingesetzt wurde [13]. Der IDOG verwendet das in [3] beschriebene Navigationsframework. Physikalische Inspektionsposenmarker werden an den entsprechenden Inspektionsposen im realen Gebäude platziert. Es werden drei Inspektionseinsätze durchgeführt, um die Genauigkeit anhand der Abweichung zwischen den geplanten Inspektionsposen und den tatsächlich erreichten Posen des IDOG zu testen. Für die Zuverlässigkeit des BIM-basierten Inspektionsframeworks werden die Anzahl der manuellen Eingriffe aufgezeichnet. Die Ergebnisse werden im folgenden Unterabschnitt zusammengefasst.

3.3 Ergebnisse

Alle fünf Inspektionsposen werden erfolgreich erzeugt. Die Ausrichtung des BIM-Modells und der Karte für die Roboternavigation wird mit geringen Abweichungen durchgeführt. Die Ergebnisse der Validierungstests hinsichtlich Genauigkeit sind in Tabelle 1 dargestellt. Im RMSE zusammengefasst, wird ein Fehler von 0,129 m beobachtet. Der IDOG navigiert zu den Inspektionsposen ohne manuellen Eingriff (Abbildung 5). Das Inspektionsmodul ist in der Lage, die BT-Strukturen zuverlässig auszuführen und die Zielelemente, d. h. die Steckdosen, zu erkennen und den Zustand erfolgreich durch die Aufnahme von Bildern inklusive Zeitstempel zu dokumentieren.

Tabelle 1: Abweichungen zwischen geplanten und tatsächlichen Posen des IDOGs in Metern

Test Nr.	Pos 1	Pos 2	Pos 3	Pos 4	Pos 5
1	0,18	0,12	0,18	0,14	0,11
2	0,09	0,06	0,16	0,07	0,22
3	0,13	0,06	0,13	0,07	0,07

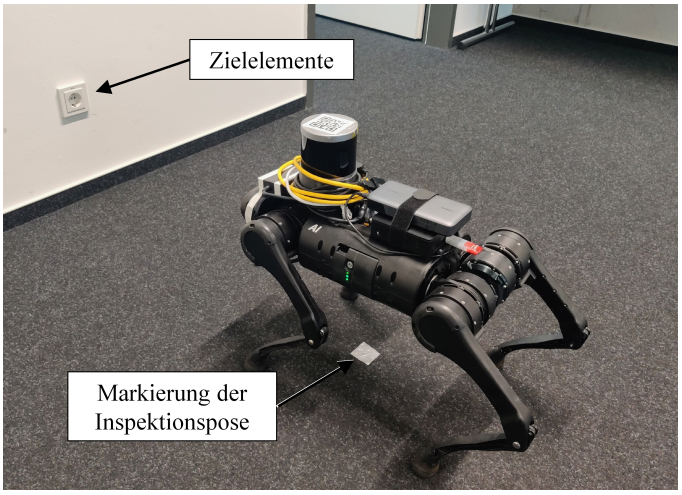


Abbildung 4: Vierbeiniger Roboter "IDOG" in den Validierungstests

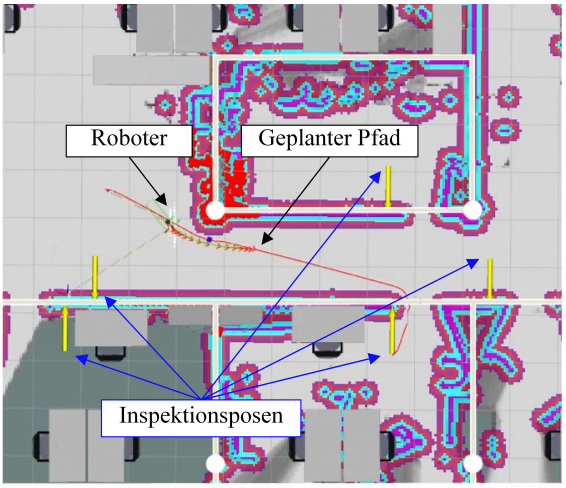


Abbildung 5: IDOG navigiert zu Inspektionsposen

3.4 Diskussion

Die Abweichungen zwischen den geplanten und den vom IDOG erreichten Inspektionsposen werden von der Genauigkeit der Platzierung der AprilTags und der Zielelemente im BIM-Modell beeinflusst. Abweichungen zwischen realen und der im BIM-Modell platzierten virtuellen AprilTags führen zu Fehlausrichtungen der Karte und des BIM-Modells. Abweichungen zwischen den Posen der realen und virtuellen Zielelemente im BIM-Modell führen zu Messfehlern. Insgesamt ist die Genauigkeit mit einem RMSE von 0,129 m für visuelle Inspektionen ausreichend. Das BIM-basierte Inspektionsframework hat sich als zuverlässig erwiesen, da alle Inspektionsaufträge von den IDOGs ohne menschliches Eingreifen ausgeführt wurden. Im nächsten Abschnitt wird die Arbeit zusammengefasst und es werden Schlussfolgerungen gezogen.

4 Zusammenfassung und Schlussfolgerungen

BIM-Modelle enthalten wertvolle semantische und geometrische Informationen, um automatisierte Gebäudeinspektionen mit mobilen Robotern zu verbessern. In diesem Beitrag wurden die Herausforderungen der Interoperabilität angegangen, indem künstliche Referenzmarken zur Kartenausrichtung sowie BT-Kontrollstrukturen eingesetzt wurden. Das entwickelte BIM-basierte Inspektionsframework wurde per automatischer visueller Inspektion von Steckdosen mit einem vierbeinigen Roboter in einer Büroumgebung validiert. Die Validierungstests zeigen die Fähigkeit des Systems, genaue und zuverlässige Gebäudeinspektionen zu ermöglichen. Die Extraktion von Gebäudedaten aus BIM-Modellen ermöglichte eine kontextbewusste Navigation, die Integration von künstlichen Referenzmarken ermöglichte eine genaue Kartenausrichtung und die BT-Kontrollstrukturen boten einen

robusten Mechanismus zur Anpassung von Inspektionsaufgaben an reale Umgebungen. In zukünftigen Arbeiten könnte das Framework erweitert werden, um verschiedene Arten von Gebäudeelementen und Inspektionsaufgaben zu integrieren. Weiterhin könnte mehreren Robotern die gemeinsame Durchführung von Inspektionsaufgaben ermöglicht werden.

Literatur

[1] S. Halder und K. Afsari, »Robots in inspection and monitoring of buildings and infrastructure: A systematic review«, *Applied Sciences*, Jg. 13, Nr. 4, S. 2304, 2023.

[2] M. H. Shariq und B. R. Hughes, »Revolutionising building inspection techniques to meet large-scale energy demands: A review of the state-of-the-art«, *Renewable and Sustainable Energy Reviews*, Jg. 130, S. 109 979, 2020.

[3] A. Tandon, J. Stührenberg, K. Dragos und K. Smarsly, »Autonomous navigation of quadruped robots for monitoring and inspection of civil infrastructure«, in *Advances in Information Technology in Civil and Building Engineering*, A. Francis, E. Miresco und S. Melhado, Hrsg., Bd. 629, Cham: Springer Nature Switzerland, 2025, S. 347–360.

[4] K. Smarsly, K. Dragos, J. Stührenberg und M. Worm, »Mobile structural health monitoring based on legged robots«, *Infrastructures*, Jg. 8, Nr. 9, S. 136, 2023.

[5] M. Theiler und K. Smarsly, »IFC Monitor – An IFC schema extension for modeling structural health monitoring systems«, *Advanced Engineering Informatics*, Jg. 37, S. 54–65, 2018.

[6] J. Stührenberg und K. Smarsly, »LIO-BIM – Coupling lidar inertial odometry with building information modeling for robot localization and mapping«, *Advanced Engineering Informatics*, Jg. 66, S. 103 477, 2025.

[7] K. Kim und M. Peavy, »BIM-based semantic building world modeling for robot task planning and execution in built environments«, *Automation in Construction*, Jg. 138, S. 104 247, 2022.

[8] J. Chen, W. Lu, Y. Fu und Z. Dong, »Automated facility inspection using robotics and BIM: A knowledge-driven approach«, *Advanced Engineering Informatics*, Jg. 55, S. 101 838, 2023.

[9] A. Tandon, J. Stührenberg und K. Smarsly, »Automated building inspections coupling building information modeling and behavior trees«, in *Proceedings of the 2025 European Conference on Computing in Construction (EC3)*, Porto, Portugal, Juli 2025.

[10] R. Ghzouli, T. Berger, E. B. Johnsen, A. Wasowski und S. Dragule, »Behavior trees and state machines in robotics applications«, *IEEE Transactions on Software Engineering*, Jg. 49, Nr. 9, S. 4243–4267, 2023.

[11] E. Olson, »AprilTag: A robust and flexible visual fiducial system«, in *2011 IEEE international conference on robotics and automation*, IEEE, 2011, S. 3400–3407.

[12] G. Heppner, N. Berg, D. Oberacker, N. Spielbauer, A. Roennau und R. Dillmann, »Distributed behavior trees for heterogeneous robot teams«, in *2023 IEEE 19th International Conference on Automation Science and Engineering (CASE)*, 2023, S. 1–8.

[13] S. Johann u. a., »Implementation and validation of robot-enabled embedded sensors for structural health monitoring«, *The e-journal of nondestructive testing & ultrasonics*, 2024.

Von Materialtests zum automatisierten Monitoring von 3D-gedruckten Lehmbauwerken

Jasper Vollmert, P. Peralta and A. Alatassi
Institut für Digitales und Autonomes Bauen, Technische Universität Hamburg, Blohmstraße 15, 21079 Hamburg
E-mail(s): jasper.vollmert@tuhh.de, patricia.peralta.abadia@tuhh.de, adel.alatassi@tuhh.de

Abstract: Das Bauwerksmonitoring (engl. „Structural Health Monitoring“, SHM) ist eine bewährte Praxis zur Gewährleistung der Sicherheit und Zuverlässigkeit von Bauwerken. Mit der steigenden Nachfrage nach umweltverträglichen Bauverfahren und der Notwendigkeit, den CO_2-Fußabdruck von Bauprojekten zu reduzieren, gewinnen nachhaltige Materialien wie Lehm an Bedeutung. Lehmgedruckte Strukturen führen zu einem neuen Bereich des Bauwerksmonitorings, der Anpassungen etablierter SHM-Strategien erfordert. Die Forschung zu SHM-Strategien für lehmgedruckte Strukturen ist spärlich, was eine Lücke im Verständnis der langfristigen Leistungsfähigkeit lehmgedruckter Strukturen hinterlässt. Diese Arbeit entwickelt eine Methodik zur experimentellen Untersuchung des strukturellen Verhaltens lehmgedruckter Bauteile – einschließlich Beulverhalten, Schrumpfung und Tragfähigkeit – und leitet daraus zentrale Faktoren für SHM ab. Die Methodik umfasst Zustandsbewertung, Randbedingungen, Designoptimierung, Prototyping und SHM-Strategiedefinition. Anhand eines Wandbauteils wird die Schrumpfung untersucht und eine SHM-Strategie definiert, bestehend aus Sensorwahl, Sensorplatzierung und Festlegung von Entscheidungskriterien. Die Ergebnisse zeigen die Produzierbarkeit stabiler Lehmstrukturen und liefern wertvolle Impulse für nachhaltiges Bauen.

Keywords: Additive Fertigung, Lehmdruck, Bauwerksmonitoring, nachhaltiges Bauen

DOI: 10.18154/RWTH-CONV-254869. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Einleitung

Digitale Fertigungstechniken, insbesondere der 3D-Druck (engl. „3D Printing“, 3DP), haben in den letzten Jahren in verschiedenen Branchen zunehmend Verbreitung gefunden [1]. Der Trend zur Nutzung von 3DP im Bauwesen hat die Forschung zur Automatisierung von Bauprozessen gefördert, aufgrund gut entwickelter digitaler Baumethoden sowie der Forschung zu 3DP-Materialien wie Beton und Lehm [2]. Daten zur langfristigen Leistungsfähigkeit von 3DP-Strukturen und zum inneren Zustand von 3DP-Materialien sind wesentlich für das Verständnis der Beziehungen zwischen 3DP-Prozessen, Materialien und Geometrien [3]. Durch das Bauwerksmonitoring (engl. „Structural Health Monitoring“, SHM), also der Erfassung und Bewertung struktureller Merkmale, wie Bucklungsverhalten, Schrumpfung und Tragfähigkeit, lässt sich die langfristige Leistungsfähigkeit von 3D-gedruckten Strukturen beurteilen. Allerdings fehlen bislang SHM-Methoden, die speziell auf die Besonderheiten von 3D-gedruckten

Lehmbauten zugeschnitten sind [4].
Im Allgemeinen können eingebettete Sensoren das Monitoring des strukturellen und materiellen Verhaltens von lehmgedruckten Strukturen verbessern [5]. In 3DP-Anwendungen wurden eingebettete Sensoren hauptsächlich für den Betondruck und in geringerem Maße für den Lehmdruck eingesetzt [6]. Da die Sensorik im Betondruck im Vergleich zur Sensorik im Lehmdruck ausgereifter ist [7], können Synergien zwischen Methodiken im Beton- und Lehmdruck genutzt werden, um strukturelle Parameter, wie Dehnung und Verformung, sowie Materialparameter, wie Temperatur und Feuchtigkeitsgehalt, von lehmgedruckten Strukturen zu überwachen [8].
In dieser Arbeit wird eine Methodik zur Definition einer SHM-Strategie für lehmgedruckte Strukturen vorgeschlagen. Die Methodik berücksichtigt strukturelles und materielles Verhalten und umfasst Zustandsbewertung, Randbedingungsdefinition, Designoptimierung, Prototyping und schließlich SHM-Strategiedefinition. Der Rest dieser Arbeit ist wie folgt strukturiert: Zunächst wird ein Überblick über die vorgeschlagene Methodik gegeben, diese experimentell getestet, um eine SHM-Strategie abzuleiten, anschließend die Wirksamkeit der SHM-Strategie durch Einbettung von Sensoren zum Monitoring des Feuchtigkeitsgehalts validiert. Abschließend wird mit einer Zusammenfassung der wichtigsten Erkenntnisse geschlossen.

2 Methodik und experimentelle Tests

Die Methodik besteht aus fünf Phasen: Zustandsbewertung, Randbedingungsdefinition, Designoptimierung, Prototyping und SHM-Strategiedefinition (Abbildung 1). In der ersten Phase werden die Ausgangsbedingungen festgelegt, um Klimabedingungen, Materialeigenschaften, verfügbare Geräte und Softwareanwendungen zu identifizieren. Lehmmischungen werden durch Materialtests auf Extrudierbarkeit, Baubarkeit, kleinskalige Schrumpfung und Rissbildungsverhalten bewertet. In der zweiten Phase werden Randbedingungen durch strukturelle, fertigungstechnische und sensorische Einschränkungen definiert, wobei fertigungstechnische Randbedingungen durch Prozessparametertests bestimmt werden [9]. Die dritte Phase umfasst die Designoptimierung, bei der Designvariablen identifiziert und auf Grundlage der definierten Randbedingungen optimiert werden, wobei parametrisches Design für Füllmuster und Verbindungen genutzt wird und Softwareanwendungen wie Rhino 8 [10] mit Grasshopper 3D verwendet werden. In der vierten Phase werden physische Prototypen entwickelt und evaluiert, um das Design zu verfeinern und strukturelle Integrität sowie Herstellbarkeit zu gewährleisten, wobei ein Prototyp pro Lehmmischung produziert wird. Die fünfte Phase definiert die SHM-Strategie basierend auf dem experimentell bewerteten Strukturverhalten, berücksichtigt Wechselwirkungen zwischen Strukturverhalten und parametrischem Design bei der Definition von Sensortyp und -Positionen und bestimmt Sensorpositionen anhand potenzieller Ausfallszenarien. Eine Rückkopplungsschleife erleichtert die Anpassung des parametrischen Designs an die definierte SHM-Strategie, wobei Sensoren in lehmgedruckte Prototypen eingebettet werden, um die SHM-Strategie zu validieren. Die Methodik wird durch experimentelle Tests implementiert, um die SHM-Strategie abzuleiten. In dieser Arbeit liegt der Schwerpunkt auf den experimentellen Phasen des Prototyping und der SHM-Strategiedefinition (Phasen 4-5), während die Phasen 1-3 im Folgenden nicht weiter vertieft werden.

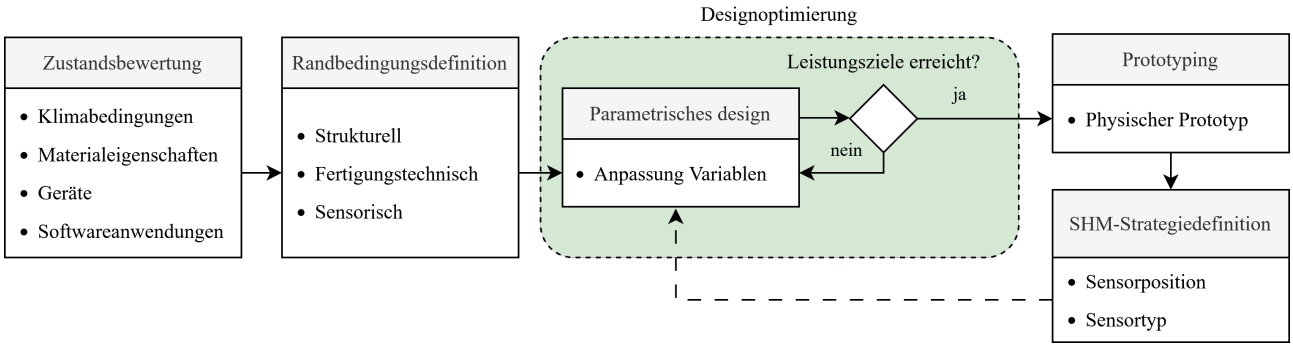


Abbildung 1: Allgemeiner Arbeitsablauf zur Festlegung von SHM-Strategien.

2.1 Prototyping

Zwei Prototypenexemplare einer Wandkomponente, Prototyp A (bestehend aus Lehm A) und Prototyp B (bestehend aus Lehm B), werden gedruckt und auf Beulverhalten, Überhangwinkel und Oberflächenqualität unmittelbar nach der Herstellung sowie auf Schrumpfung und Rissbildung während und nach dem Trocknungsprozess bewertet. Die Prototypen werden als druckbar angesehen, ohne Ausfälle oder signifikantes Beulen, zeigen jedoch nicht einheitliche Trocknungsmuster mit einem gerichteten Gradienten von oben nach unten und von äußeren Oberflächen zur inneren Füllung. Die Feuchtigkeitsgehaltsbewertung nach sechstägiger Trocknungsperiode zeigt Unterschiede zwischen den Prototypen: Prototyp A demonstriert eine schnellere Trocknungsrate mit schnellerer Feuchtigkeitsabgabe, während Prototyp B eine langsamere Trocknungsrate aufweist, erkennbar durch dunklere Verfärbung der Füllung und Konturen, was auf zurückgehaltene Feuchtigkeit hinweist.

Trocknungsrisse werden während des Trocknungsprozesses dokumentiert. Prototyp A entwickelt mehrere Trocknungsrisse (Abbildung 2), wobei die ersten Risse nach Tag vier in der Mittellinie-Region erscheinen und von unten nach oben fortschreiten. Weitere Risse entstehen an den Tagen sechs und acht in der Mittellinie und vorderen Region. Der Hauptriss in der Mittellinie-Region stellt aufgrund seiner Lage und Größe einen potenziellen Schwachpunkt dar. Prototyp B zeigt deutlich reduzierte Trocknungsrisse mit nur einem einzigen Riss nach acht Tagen in der Mittellinie-Region, der ähnlich von unten nach oben fortschreitet, jedoch mit geringerer Schwere.

Nach dem Trocknen zeigen die Prototypen ungleichmäßige Schrumpfung: Prototyp A weist eine geringere horizontale Schrumpfung als Prototyp B. Das unterschiedliche Schrumpfungsverhalten resultiert aus der Trocknungsrate und Druckfehlern, wobei erhöhter Wassergehalt zu verstärkter Schrumpfung führt und Druckfehler zusätzliche Rissbildung durch innere Spannungen verursachen können. Für die SHM-Strategie wird eine zweite Prototypiteration mit eingebetteten Feuchtigkeitssensoren durchgeführt, wobei die Mischung von Lehm B für Prototyp B2 gewählt wird, um die Genauigkeit des SHM-Systems zu bewerten und die Sensorpositionierung zu optimieren.

2.2 Sensorintegration zur SHM-Strategiedefinition

Als Grundlage für die Definition der SHM-Strategie wird die Integration verschiedener Sensortypen und -positionen untersucht. Ein Validierungstest wird durchgeführt, um die Leistungsfähigkeit von resistiven und kapazitiven Feuchtigkeitssensoren sowie die Auswirkungen der Sensorpositionen

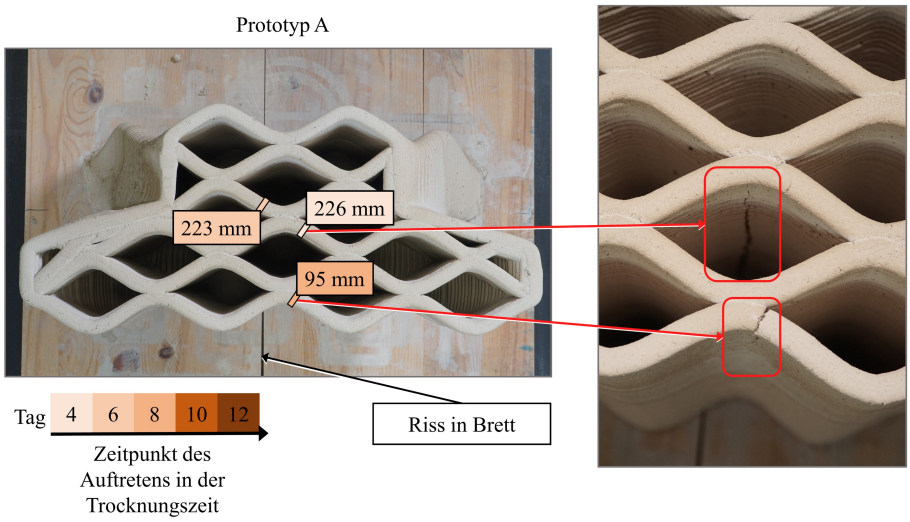


Abbildung 2: Austrocknungsrisse in Prototyp A.

zu bewerten. Resistive und kapazitive Feuchtigkeitssensoren werden typischerweise zur Messung des Bodenfeuchtigkeitsgehalts mittels verschiedener Messprinzipien verwendet, wobei der Einsatz beider Sensortypen zum Monitoring von Trocknungsprozessen im Lehmdruck einen Vergleich der Empfindlichkeit für die Erkennung von Feuchtigkeitsänderungen ermöglicht. Wie in Abbildung 3 dargestellt, werden zwei Sensorsysteme zum Monitoring des Prototyps B2 verwendet: Sensorsystem I, ein kommerzielles System mit einem resistiven Feuchtigkeitssensor zum Bodenmonitoring, und Sensorsystem II, ein maßgeschneidertes SHM-System mit einem Feuchtigkeitssensorknoten (drei kapazitive Feuchtigkeitssensoren) und einem Umgebungssensorknoten zum kontinuierlichen Monitoring des Lehmfeuchtigkeitsgehalts und der Umgebungsbedingungen.

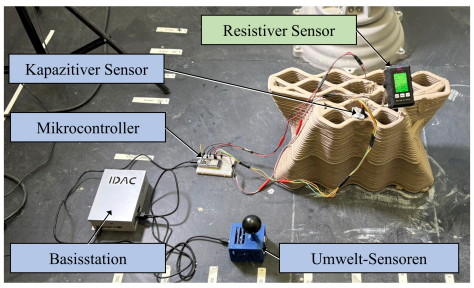


Abbildung 3: Komponenten des Sensorsystems I (blau) und des Sensorsystems II (grün)

Die resistiven und kapazitiven Feuchtigkeitssensoren werden in verschiedenen Einbettungstiefen positioniert und in äußeren und inneren Lagen des Prototyps B2 verteilt, um den Trocknungsprozess gemäß dem beobachteten Trocknungsgradienten zu überwachen. Sensorsystem I ist für Bodenfeuchtigkeit durch Widerstandswerte zwischen zwei Sonden ausgelegt, während Sensorsystem II gleichzeitig für Echtzeitdatenaufzeichnung des kapazitiven Feuchtigkeitsgehalts und der Umgebungsbedingungen

konzipiert ist. Der Feuchtigkeitssensorknoten umfasst drei eingebettete kapazitive Sensoren.
Die Positionierung wird durch geometrische Einschränkungen bestimmt: der resistive Sensor besteht aus zwei 180 mm langen Sonden im Abstand von 30 mm, während die kapazitiven Sensoren 98×23×4 mm messen. Bei 11 mm Schichtbreite werden die Sensoren vertikal an überlappenden Punkten eingebettet, um strukturelle Stabilität zu gewährleisten. Die kapazitiven Sensoren werden gemäß Druckprozess positioniert: Sensor 1 und Sensor 2 vollständig eingebettet, Sensor 3 und der resistive Sensor 4 teilweise eingebettet. Prototyp B2 wird über 10 Tage beobachtet ohne strukturelle Instabilitäten oder sensorinduzierte Risse. Die Umgebungsdaten zeigen Feuchtigkeit zwischen 15-40% und Temperaturen zwischen 18,1-23,8°C, wobei der abnehmende Feuchtigkeitsgehalt in der Struktur diesen Umgebungsänderungen entspricht (Abbildung 4). Kapazitive Sensoren zeigen eine 25-stündige Stabilisierungsphase, während vollständig eingebettete Sensoren nahezu identische Trends aufweisen und der teilweise eingebettete Sensor 3 in den ersten 160 Stunden höhere anfängliche Feuchtigkeitswerte zeigt.
Darüber hinaus zeigt, wie aus Abbildung 4 ersichtlich, der resistive Sensor (Sensor 4) im Vergleich zu den kapazitiven Sensoren in den ersten 195 Stunden der Messung höhere Feuchtigkeitswerte, gefolgt von einem starken Abfall des Feuchtigkeitsgehalts. Obwohl ein ähnlicher Feuchtigkeitsgehaltstrend zwischen Sensor 2 und Sensor 4 erwartet wurde, deuten die für Sensor 4 beobachteten Werte darauf hin, dass das widerstandsbasierte Messprinzip im Vergleich zum kapazitiven Messprinzip nicht so empfindlich gegenüber der Änderung des Feuchtigkeitsgehalts ist. Durch Beobachtung des kumulativen Bodenfeuchtigkeitsverlustes im Zeitverlauf (Abbildung 5) wird ein Gesamtfeuchtigkeitsverlust von 43% bis 48% beobachtet. Die gemessenen kumulativen Feuchtigkeitsverluste von Sensor 1 und Sensor 3 sind vergleichbar, während Sensor 2 leicht in einem akzeptablen Bereich abweicht. Im Gegensatz dazu weicht Sensor 4 erheblich ab, was auf eine begrenzte Genauigkeit und Ansprechbarkeit hinweist.

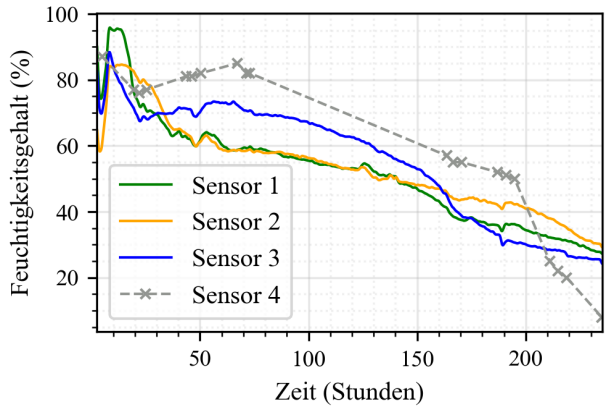


Abbildung 4: Messwerte der Feuchtesensoren.

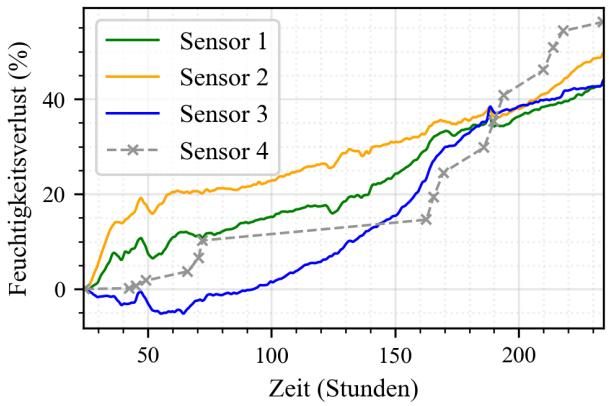


Abbildung 5: Kumulativer Feuchtigkeitsverlust über die Zeit (beginnend mit t = 25 h).

3 Diskussion und Definition der SHM-Strategie

3.1 Allgemeine Diskussion der Ergebnisse

Die Feuchtigkeitssensoren haben unterschiedliche Genauigkeiten gezeigt. Vollständig eingebettete Sensoren haben nahezu identische Werte gezeigt, während die teilweise eingebetteten Sensoren

tendenziell höhere Feuchtigkeitswerte aufgezeichnet haben. Nach anfänglichem Oberflächenwasserverlust haben sich die Sensorwerte mit den visuellen Beobachtungen in Einklang gebracht und reflektieren die Trocknungsbedingungen des Prototyps B2. Während die absoluten Feuchtigkeitsmessungen zwischen den Sensortypen inkonsistent waren, haben die aufgezeichneten Trends der kapazitiven Bodenfeuchtigkeitssensoren mit den visuellen Bewertungen korreliert, was die Eignung kapazitiver Messungen für das Monitoring nahelegt. Die durchlässigen Membranen haben die Feuchtigkeitspenetration verzögert und zusätzliche Zeit für die Sensorstabilisierung verursacht. Kapazitive Änderungen wurden von den eingebetteten Sensoren mit unterschiedlicher Genauigkeit erfasst, abhängig von Sensortyp und Position; die Messungen der kapazitiven Feuchtigkeitssensoren haben jedoch erfolgreich den Trocknungsgradienten innerhalb des Prototyps B2 erfasst. Die Umgebungsbedingungen haben den Trocknungsprozess erheblich beeinflusst, wobei ein beschleunigter Feuchtigkeitsverlust bei niedrigerer Luftfeuchtigkeit und erhöhter Temperatur beobachtet wurde. Insgesamt unterstreichen die Ergebnisse die Anwendbarkeit von Feuchtigkeitssensoren zum Monitoring des Feuchtigkeitsverlusts in additiv gefertigten Lehmstrukturen und bilden die Grundlage für die Entwicklung einer maßgeschneiderten SHM-Strategie.

3.2 SHM-Strategie

Die Definition einer SHM-Strategie für lehmgedruckte Strukturen erfordert eine gezielte Sensorauswahl zum Monitoring von Lehmfeuchtigkeit und struktureller Integrität. Kapazitive Feuchtigkeitssensoren erwiesen sich als zuverlässiger und empfindlicher als resistive Sensoren, die anfälliger für Korrosion sind. Die Sensorplatzierung sollte den Trocknungsgradienten folgen, um Feuchtigkeitsmigration zu verfolgen und rissanfällige Zonen zu identifizieren.

Ein zentrales Problem stellt der Einbettungszeitpunkt dar. Kabelgebundene Sensoren bergen Rissbildungsrisiken an Übergangspunkten durch Lehmschrumpfung und sind aufgrund der geometrischen Komplexität 3D-gedruckter Komponenten schwer skalierbar. Drahtlose Lösungen sind daher unerlässlich, jedoch stellen die hohe Dielektrizitätskonstante und Feuchtigkeitsretention von Lehm erhebliche Herausforderungen für die Signalübertragung dar. Erhöhte Feuchtigkeit verstärkt die Signaldämpfung durch gesteigerte Materialkonduktivität, während Mehrwegeinterferenz zu Signalverzerrung führen kann.

Für zuverlässige drahtlose Sensorik sind adaptive Kommunikationsprotokolle und sorgfältige Frequenzauswahl erforderlich, die den variablen elektromagnetischen Eigenschaften während des Trocknungsprozesses Rechnung tragen. Aktuelle technologische Limitationen umfassen unzureichende Genauigkeit kostengünstiger kapazitiver Sensoren, Korrosionsanfälligkeit von Widerstandssensoren und begrenzte Empfindlichkeit gegenüber graduellen Feuchtigkeitsänderungen. Die Implementierung erfolgreicher SHM-Strategien erfordert daher die Entwicklung robuster, miniaturisierter drahtloser Sensoren mit verbesserter Messgenauigkeit für langfristige Echtzeitdatenerfassung in der anspruchsvollen Umgebung lehmgedruckter Strukturen.

4 Zusammenfassung und Schlussfolgerungen

Mit der zunehmenden Aufmerksamkeit für Lehmdruck in Forschung und Praxis müssen SHM-Strategien an Lehmdrucktechnologien angepasst werden. In dieser Studie wurde eine SHM-Strategie vorgeschlagen, und die strukturellen Eigenschaften von lehmgedruckten Strukturen, einschließlich Beulen, Schrumpfung und Tragfähigkeit, wurden experimentell bestimmt. Prototyping und Validierung der Sensorplatzierung haben die Machbarkeit der Konstruktion strukturell stabiler lehmgedruckter Strukturen demonstriert und wichtige Einblicke in das Materialverhalten (d.h. Schrumpfung) für die Ableitung von Strategien zur Sensorpositionierung geliefert.

Die in dieser Studie vorgeschlagene Methodik hat einen strukturierten Ansatz zum Verständnis von Lehm als 3DP-Material geboten, der sowohl die Materialcharakterisierung als auch eine SHM-Strategie umfasst. Die experimentellen Ergebnisse haben gezeigt, dass eine effektive Sensorplatzierung durch systematische Analyse des Materialverhaltens erreicht werden kann, wobei der Fokus insbesondere auf Trocknungsmustern und potenziellen Fehlerstellen liegt. Darüber hinaus hat die Studie die Beziehungen zwischen kleinmaßstäblichen Labortests und größeren Prototypen behandelt und wesentliche Skalierbarkeitsüberlegungen für SHM-Sensornetzwerke großer Strukturen aufgestellt. Zusammenfassend tragen die Ergebnisse zur Weiterentwicklung des Bereichs nachhaltigen Bauens bei, indem sie die Lebensfähigkeit des Lehmdrucks auf der Grundlage einer fundierten SHM-Strategie demonstrieren. Zukünftige Forschungen könnten die Erweiterung der SHM-Strategie von kabelbasierten zu drahtlosen Systemen umfassen.

Acknowledgements

Die Autoren bedanken sich für die finanzielle Unterstützung durch die Deutsche Forschungsgemeinschaft (DFG) im Rahmen der Projekte SM 281/22-1 und SM 281/31-1 sowie durch die Technische Universität Hamburg über das I^3-Programm im Rahmen des Nachwuchsprojekts „Digitalization of Earth Printing to Advance Climate-Informed Engineering“. Die Autoren danken außerdem der Goerg & Schneider GmbH & Co. KG für die Bereitstellung des Tonmaterials. Die in diesem Beitrag dargestellten Meinungen, Erkenntnisse, Schlussfolgerungen und Empfehlungen stammen ausschließlich von den Autoren und spiegeln nicht zwangsläufig die Ansichten der fördernden Organisationen wider.

Literatur

[1] A. Kazemian, X. Yuan, E. Cochran und B. Khoshnevis, »Cementitious materials for construction-scale 3D printing: Laboratory testing of fresh printing mixture«, *Construction and Building Materials*, Jg. 145, S. 639–647, 2017.

[2] P. Peralta und K. Smarsly, »Requirements analysis of additive manufacturing for concrete printing - A systematic review«, in *Proceedings of the 39th International Symposium on Automation and Robotics in Construction*, Bogota, Colombia, Juli 2022.

[3] K. Smarsly, P. Peralta, D. Luckey, S. Heine und H.-M. Ludwig, »BIM-based concrete printing«, in *Proceedings of the International ICCCBE and CIB W78 Joint Conference on Computing in Civil and Building Engineering 2020*, Sao Paolo, Brazil, Aug. 2020.

[4] P. Peralta, M. E. Ahmad und K. Smarsly, »Printing Information Modeling (PIM) for Additive Manufacturing of Concrete Structures«, *Applied Sciences*, Jg. 13, Nr. 23, 2023.

[5] G. Ma, Y. Li, L. Wang, J. Zhang und L. Li, »Real-time quantification of fresh and hardened mechanical property for 3D printing material by intellectualization with piezoelectric transducers«, *Construction and Building Materials*, Jg. 241, 2020.

[6] C. Hong, Y. Zhang und L. Borana, »Performance investigation of 3D printed clay soil using fiber Bragg grating technology«, *Acta Geotechnica*, Jg. 17, Nr. 2, S. 453–462, 2022.

[7] P. Peralta, S. Heine, H.-M. Ludwig und K. Smarsly, »A BIM-based approach towards additive manufacturing of concrete structures«, in *Proceedings of the 27th International Workshop on Intelligent Computing in Engineering*, Berlin, Germany, Juli 2020.

[8] A. Perrot, D. Rangeard und T. Lecompte, »Field-oriented tests to evaluate the workability of cob and adobe«, *Materials and Structures*, Jg. 51, 2018.

[9] A. Curth, N. Pearl, A. Castro-Salazar, C. Mueller und L. Sass, »3D printing earth: Local, circular material processing, fabrication methods, and life cycle assessment«, *Construction and Building Materials*, Jg. 421, 2024.

[10] Robert McNeel & Associates, *Rhinoceros 3D*, https://www.rhino3d.com.

Mobile Application to Support Facade Renovation

Oliver Quintel[1,2], Hendrik Martin Würz[1,2], Michel Krämer[1,2], Eva Klien[1] and Arjan Kuijper[1,2]
[1]Fraunhofer Institute for Computer Graphics Research IGD, Darmstadt, Germany
[2]Technical University of Darmstadt, Germany
E-mail(s): oliver@quintel.net, hendrik.martin.wuerz@igd.fraunhofer.de

Abstract: We present a mobile application that enables end users to quickly calculate the size of a building facade. Our goal is to support the renovation of facades, as their costs are mainly determined by the surface area. We developed an image-based approach that consists of four steps. First, we measure the facade width using augmented reality (AR) functionalities of modern smartphones. In our evaluation, this measurement achieves an average accuracy of 98.5 %. In the second step, the user takes a photo of the facade and marks its contours. All information is transferred to a server that rectifies the photo. The result of step three is a distortion-free orthophoto of the facade, with the correct ratio of width to height. The fourth and final step is AI-based image recognition to identify facade elements such as windows, doors, and roofs. The facade area is calculated by subtracting the area of facade elements from the total area. We evaluated our application on ten different buildings and achieved an average accuracy of 97.5 % for the facade area.

Keywords: Facade renovation, Mobile applications, Augmented reality (AR)

DOI: 10.18154/RWTH-CONV-254870. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Climate change continues to progress and requires consistent action to reduce greenhouse gas emissions. The building sector plays a key role, as it is a major contributor to emissions due to its high energy demand for heating and electricity. Older buildings in particular lose significant amounts of energy due to poor insulation. Energy-efficient renovation of the facades is therefore urgently needed, but the implementation of renovation measures still depends heavily on the motivation of homeowners. They often do not know how much a renovation will cost and are overwhelmed by the complexity of the process. There are online calculators providing cost estimates, but they need the area of the facade as input, which is often unknown to homeowners. A simple and easy-to-use tool for calculating the facade area has the potential to significantly reduce the entry barrier to the renovation process.

In this paper, we present such a tool for modern smartphones. We developed an image-based approach since it is easy to use and does not require any special hardware. We use AI-based image recognition to calculate the facade area in pixels. To be able to convert this information to square meters, we need a reference size from the real world. Our app uses Augmented Reality (AR) functionalities of modern smartphones to measure the facade width as a reference size.

2 Related Work

There are a few existing methods to calculate the area of a building facade. First, it is of course always possible to use a measuring tape. This approach is simple but requires time. It can even become a dangerous endeavor if a building is very high. Laser measuring devices provide a safe and accurate alternative, but they are expensive and not applicable to complex surfaces.

Professional solutions are based on 3D point clouds, which are recorded using LiDAR devices (Light Detection and Ranging). Point clouds have been used to reconstruct building facades [1]. The results are very accurate, but problems can arise if the laser hits transparent surfaces, which can lead to gaps and other measuring errors [2]. Also, scanning a building facade with such a professional device is time-consuming and expensive. There is also a lack of automated data processing [1]. Modern AI-based processing approaches are promising, but still require manual effort [3].

Another alternative is provided by close-range photogrammetry [4]. Here, multiple images are taken around an object with convergent camera orientations. Through mathematical techniques such as image registration and bundle adjustment, a 3D model can be derived, which in turn allows the facade to be measured. However, this approach requires planning effort, careful acquisition, sometimes even sophisticated gear and setups [5], as well as complex post-processing. Similar to laser scanning, it is therefore not suitable for private homeowners who simply want to know the area of their facade.

For this reason, we present a low-cost app-based solution in this paper. Our app uses augmented reality (AR), which has already been recognized as a valuable technology in the building construction sector [6]. The ARKit and ARCore frameworks for the iPhone and Android phones, respectively, provide means for developers to integrate AR functionality in their mobile apps [7]. Maneli and Isafiade have evaluated these frameworks with respect to the measuring quality [8], [9]. They report accuracies of 89.42 % (ARCore) and 97.52 % (ARKit), increasing to 99.63 % (ARCore) and 99.26 % (ARKit) depending on the distance, which is more than suitable for the purpose of this work.

3 Approach

With our approach we want to create both an easy-to-use application with low hardware requirements and sufficient accuracy for cost estimation. To calculate the facade area we use four steps:

1. Measuring the facade width
2. Take a photo of the facade and define its contour
3. Rectifying the photo of the facade
4. Identify facade elements

Steps 1 and 2 are realised within a smartphone app, while steps 3 and 4 are carried out on a server. The results of each step are shown in Fig. 1.

3.1 Measuring the facade width

The first step is to measure the width of the facade. This enables us to convert a size in pixels into metric data later, thus we call it *reference size*. Instead of the facade width, any other object such as

(a) AR width (b) Distorted image (c) Rectified image (d) Object detection

Figure 1: Results of each step of the process. The first figure shows the AR measurement of the facade width. The second figure shows the distorted photo taken by the user with the corner points of the contour. In the third figure the photo was rectified and in the last figure, facade elements have been identified.

windows or doors could have been used as a reference, too. However, the width of the facade is an easily recognisable size and thus well suited.

For distance measurement, we rely on the augmented reality (AR) capabilities of modern smartphones. To implement the mobile application, we use the AR framework ViroReact [10], which provides the AR capabilities of the smartphone in a React Native application. The user sees a live camera feed and can put virtual objects in the world. We use this to mark the bottom left and right corners of the facade as shown in Fig. 1a. Since ViroReact provides 3D position of the objects in a metric space, we can calculate the distance between them and thus the width of the facade.

3.2 Taking a photo of the facade and defining its contour

Once the facade width has been determined, the user takes a photo of the whole facade. It is important to ensure that the entire facade is visible since missing parts will not be included in the final area calculation.

Next, the user defines the contour of the facade in the photo, as shown in Fig. 1b. Since facades can have multiple corner points, the app provides three possible shapes as illustrated in Fig. 2. The user drags the corner points of a selected shape onto the photo taken before.

Shape one in Fig. 2 defines a facade with four corner points and is a simple rectangle. It is extended in shape two by an attic storey, which leads to five corner points. These two shapes are called *simple contours* since they identify the whole facade area. Although the majority of existing facades can be described by them, we recognise that there are also more complicated facades. For these cases, we

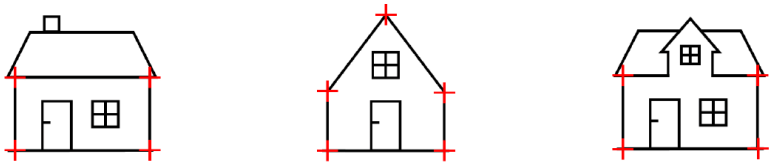


Figure 2: Different shapes for defining the facade contour.

defined a third shape, in which only a rectangular base area is selected. The upper part of the facade will be defined by an AI-based image recognition in step 4 of our process.

Instead of letting the user define the contour of the facade, we could have used a purely AI-based image recognition solution. However, our tests showed that this is prone to errors. Especially the separation of neighbouring houses or the handling of facades partially obscured by vegetation, parked cars, and other objects is too often inaccurate and did not satisfy our targeted quality.

3.3 Rectifying the photo of the facade

For the following steps, the photo, reference size, and corner points of the contour are sent from the smartphone app to a server. This reduces load on the smartphone app and improves performance.

To calculate the final facade area, we need an image in which each pixel covers the same metric area. However, the photo was taken with distortions caused by camera angle and perspective. To get a distortion free image, we use the rectification method by Zhang and He [11]. It takes four points in the distorted photo that form a rectangle in real world. In our case, we use the four lower corner points of the contour defined by the user. With these points, Zhang and He calculate the aspect ratio from width to height and return a formula for correction. We apply this to the whole photo and get a distortion-free image of the facade as shown in Fig. 1.

It has to be noted that this only applies to planar objects. Wall protrusions like bay windows are not correctly rectified and will lead to errors in the later area calculation.

3.4 Identifying facade elements

Once the image has been rectified, AI-based image recognition is used to classify the pixels. If a simple contour was defined in step 2 of our process, only the area within the contour is analysed. This ensures that no elements of neighbouring facades are taken into account. If a complex contour was selected, we additionally check the area above the upper edge. In this way, we can add additional pixels to the facade area, even if they were not included in the contour.

We use the AI framework GroundingDINO [12] to detect four classes: facade, window, door and roof. Multiple classifications of the same object are avoided by using non-maximum suppression (NMS). Within the recognized bounding boxes, we perform a segmentation to get the pixel-wise boundaries of the detected objects. For this, we use the SegmentAnything model [13]. Examples of the classification and segmentation results are shown in Fig. 1d.

We define the *total area* as the area covered by the contour, plus any facade area found above the upper edge of the complex contour. From this total area, we subtract all detected facade elements (windows, doors, roofs) to get the actual facade area. Note that this facade area is not the same as the area segmented as a facade. If cars, vegetation or other objects cover parts of the facade, they are not included in the segmentation. However, as long as they are within the contour, we still count their pixels for the facade area. This is a trade off. If, for example, a window is covered by a car, we interpret the area incorrectly as facade, but in our experience, it is much more likely that normal facade area is covered. For the same reason, we use the rectangular base area of complex contours for the total area and only rely on AI for the upper parts where occlusion is less likely.

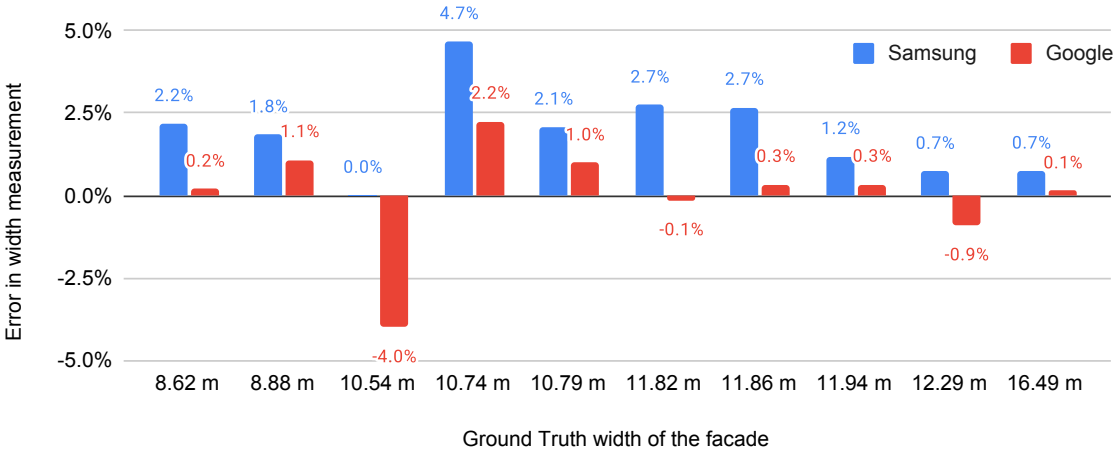


Figure 3: Error in the AR-based measurement of the facade width of the ten buildings from our study.

Finally, the facade area in pixels is converted to square meters using the reference size from step 1. The area and segmented image are sent to the app for display, allowing users to verify the process.

4 Evaluation

We evaluated the accuracy of the individual steps as well as the overall process to get clearer insights into the strengths and limitations of our approach. The app design requires on-site evaluation, so we used ten real buildings in the city of Darmstadt, Germany. They are representative for our target audience owning detached houses. For them, we had a Level of Detail (LoD) 2 model as ground truth, which was reconstructed from a point cloud.

We measured the facade width using our app on a Samsung Galaxy A32 and a Google Pixel 5 to assess hardware influence. Fig. 3 compares our results with the LoD 2 model. Both smartphones tend to overestimate width, with the Google phone being more accurate. It has an average error of 1 % (median 0.6 %, variance 1.5), while the Samsung device has 1.9 % (median 2 %, variance 1.8). This is an improvement on previous studies [8], [9], which had around 10 % error, and supports findings that devices with more powerful hardware achieve greater accuracy. However, no clear trend emerged regarding smaller or larger distances being measured more accurately.

To evaluate the rectification, we compared the height of ten buildings from the LoD 2 model with the calculated height. Since we wanted to evaluate the isolated accuracy of the rectification, we used the width of the LoD 2 model as ground truth to calculate the height. Our test buildings were between 9 and 16 m high and were reconstructed with an average error of 0.5 % (median 0.45 %, variance 0.13). In contrast to the width, the height tends to be underestimated. The intensity of the distortion had no influence on rectification. However, the more distorted the image, the more difficult it was to set the corner points in the right place. Incorrectly positioned corner points reduced the rectification quality.

To evaluate the size of the facade and its elements, we used texture information from the LoD 2 model, manually marking pixels and converting them into square meters. Due to the low visual quality of textures, object boundaries were sometimes unclear. We took this uncertainty into account by defining minimum and maximum values for each area. Overall, we got an average error of 4.1 % with the

Samsung device (median 3.1 %, variance 11.9) and 1.12 % with the Google smartphone (median 0 %, variance 6.1) The variance was noticeably higher than in other steps due to the AI-based image recognition, which relies on probabilities and is less consistent. Various error sources are discussed in Section 5. No clear trends emerged across different house categories, as both good and poor results were observed in all three.

Excess material is common in construction due to standardized packaging of materials like paint, plaster, and insulation. Our overall deviation of 2.5 % is within an acceptable range. There are no legal requirements or universal standards for material quantity calculations or cost estimate accuracy [14]. Literature suggests deviations of 30–40 % are typical for initial estimates [15], the accuracy of our application can therefore be considered accurate.

5 Discussion

In Section 4, we demonstrated that our approach enables accurate facade area calculation, though some challenges remain. The overall accuracy of the application highly depends on the accuracy of the reference size, as it determines the pixel-to-meter scale. Section 4 shows that it can be determined with a small error in general. However, occlusions from vegetation, parked cars, or poor lighting can affect its precision. Future advancements, such as LiDAR sensors in smartphones, may improve AR-based measurements.

Another challenge is the accuracy of corner points, which impacts image rectification and total area calculation. While manual corner point setting helps reduce errors from occlusions, obstacles in front of house corners can still prevent precise placement.

As noted in Section 4, image recognition can be imprecise, particularly for complex contours where additional processing is needed to determine the facade area. Fig. 4 highlights the error sources of an incorrect area calculation, marked in red. In error source 1), part of another house was mistakenly identified as facade area—an issue that occurs only with complex contours since simple

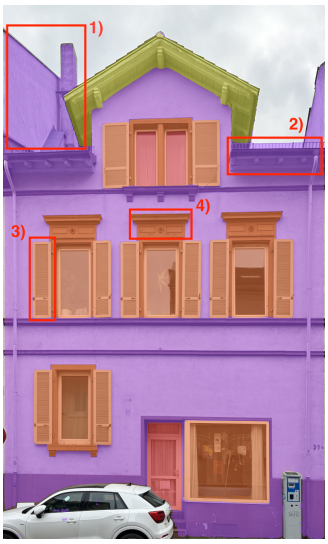


Figure 4: Example of an incorrect facade area calculation. Four different types of errors are highlighted in red.

ones completely rely on user-defined boundaries. Image recognition also struggles with certain facade elements, as shown in 2), 3), and 4). Some roof elements in 2) are not recognized as such and windows are not always detected correctly. In 3), window shutters were classified as additional window areas and as a result, too much was subtracted from the calculated facade area. Another problem is shown in 4) where window cornices are recognised as part of the windows instead of the facade. This results in the same problem as with the shutters and leads to an incorrect facade area calculation.

These problems could be avoided by using a more specific model that is trained on facade elements. However, to this date, there is no such model or dataset that is suitable for our use case [16]. Additional user input could also help improve the accuracy of the image recognition by allowing the user to correct the detected objects or add missing ones.

6 Conclusion

We presented a smartphone app that enables users to measure facade areas using augmented reality and AI-based image recognition. The application is easily accessible, does not require any additional equipment or external parameters, and is entirely based on the use of a modern smartphone.

The facade area can be calculated with an average accuracy of 97.5 %, which is within the usual margins for cost estimation. The accuracy depends on both the performance of the smartphone being used and the precision of the user interactions. The latter can be investigated in a future user study. The app is suitable for a wide range of buildings, including those with complex facades. However, the accuracy of the image recognition is crucial for the overall accuracy of the facade area calculation and can lead to significant deviations. Future work can concentrate on minimizing these errors in image recognition by using a more powerful model.

Although other methods for measuring facade areas exist, they often require expensive equipment or special knowledge. Therefore, our approach is a valuable addition to the existing methods since it can be used by anyone with a smartphone. We believe that our application has the potential to significantly reduce the entry barrier for homeowners to consider energy-efficient renovation options.

Acknowledgements

We would like to thank the city of Darmstadt for providing the LoD 2 model for our evaluation.

References

[1] Y. Xu and U. Stilla, “Toward building and civil infrastructure reconstruction from point clouds: A review on data and key techniques”, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 2857–2885, 2021. DOI: 10.1109/JSTARS.2021.3060568

[2] A. Klimkowska, S. Cavazzi, R. Leach, and S. Grebby, “Detailed three-dimensional building façade reconstruction: A review on applications, data and technologies”, *Remote Sensing*, vol. 14, no. 11, 2022. DOI: 10.3390/rs14112579

[3] J. Martínez, A. Soria-Medina, P. Arias, and A. F. Buffara-Antunes, “Automatic processing of terrestrial laser scanning data of building façades”, *Automation in Construction*, vol. 22, pp. 298–305, 2012. DOI: https://doi.org/10.1016/j.autcon.2011.09.005

[4] R. Jiang, D. V. Jáuregui, and K. R. White, “Close-range photogrammetry applications in bridge measurement: Literature review”, *Measurement*, vol. 41, no. 8, pp. 823–834, 2008. DOI: https://doi.org/10.1016/j.measurement.2007.12.005

[5] C. Ordóñez, P. Arias, J. Herráez, J. Rodríguez, and M. T. Martín, “A combined single range and single image device for low-cost measurement of building façade features”, *The Photogrammetric Record*, vol. 23, no. 122, pp. 228–240, 2008. DOI: https://doi.org/10.1111/j.1477-9730.2008.00478.x

[6] A. Z. Kolaei, E. Hedayati, M. Khanzadi, and G. G. Amiri, “Challenges and opportunities of augmented reality during the construction phase”, *Automation in Construction*, vol. 143, p. 104 586, 2022. DOI: https://doi.org/10.1016/j.autcon.2022.104586

[7] Z. Oufqir, A. El Abderrahmani, and K. Satori, “ARKit and ARCore in serve to augmented reality”, in *2020 International Conference on Intelligent Systems and Computer Vision (ISCV)*, 2020, pp. 1–7. DOI: 10.1109/ISCV49265.2020.9204243

[8] M. A. Maneli and O. E. Isafiade, “A Comparative Analysis of Augmented Reality Frameworks Aimed at Diverse Computing Applications”, in *2022 ITU Kaleidoscope- Extended reality – How to boost quality of experience and interoperability*, Dec. 2022. DOI: 10.23919/ITUK56368.2022.10003046

[9] M. A. Maneli and O. E. Isafiade, “A Multifactor Comparative Assessment of Augmented Reality Frameworks in Diverse Computing Settings”, *IEEE Access*, vol. 11, 2023. DOI: 10.1109/ACCESS.2023.3242238

[10] viro-community, *ViroReact*, en. Accessed: Dec. 4, 2024. [Online]. Available: https://viro-community.readme.io/

[11] Z. Zhang and L.-W. He, “Whiteboard scanning and image enhancement”, *Digital Signal Processing*, vol. 17, no. 2, Mar. 2007. DOI: 10.1016/j.dsp.2006.05.006

[12] S. Liu et al., “Grounding DINO: Marrying DINO with Grounded Pre-training for Open-Set Object Detection”, in *Computer Vision – ECCV 2024*, A. Leonardis, E. Ricci, S. Roth, O. Russakovsky, T. Sattler, and G. Varol, editors, Cham: Springer Nature Switzerland, 2025. DOI: 10.1007/978-3-031-72970-6_3

[13] A. Kirillov et al., “Segment Anything”, in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, Paris, France: IEEE, Oct. 2023. DOI: 10.1109/ICCV51070.2023.00371

[14] U. Fritsch, *DIN 276 in Bildern*, ger. Berlin: Beuth Verlag GmbH, 2019.

[15] H. Locher, W. Koeble, U. Locher, and W. Frick, *Kommentar zur HOAI: mit einer Einführung in das Recht der Architekten und der Ingeniere*, ger. 2002.

[16] J. Wei, Y. Hu, S. Zhang, and S. Liu, “Irregular facades: A dataset for semantic segmentation of the free facade of modern buildings”, *Buildings*, vol. 14, no. 9, Sep. 2024. DOI: https://doi.org/10.3390/buildings14092602

Structured Databases for Heritage BIM: Enhancing Digital Documentation and Restoration Workflows

Laura Marcela Valderrama Niño[1] and Sabrina Becker[1]
[1]Institute of Numerical Methods and Informatics in Civil Engineering (iib), TU Darmstadt, Darmstadt, Germany
E-mail(s): valderrama@iib.tu-darmstadt.de, becker@iib.tu-darmstadt.de

Abstract: Heritage Building Information Modeling (HBIM) plays a central role in the digital documentation and conservation of historic buildings. Yet, many current systems lack the flexibility and semantic depth required to manage evolving, multidisciplinary datasets. Rigid data structures and tool-dependent workflows hinder the representation of restoration processes, the integration of contextual knowledge, and long-term data reuse. This research introduces a hybrid data architecture that combines flexible document storage with semantic graph modeling. The system enables the integration and structured analysis of diverse heritage data—linking historical, material, and procedural information across the HBIM lifecycle. While this work is conceptual in scope, it outlines a pathway toward a modular, extensible, and open HBIM infrastructure. Aligned with ontologies such as CIDOC CRM and FAIR data principles, the proposed framework supports future integration with linked open data platforms and external heritage repositories. By advancing a semantically expressive and interoperable foundation, this approach promotes sustainable documentation, transparent conservation workflows, and deeper historical insight in digital heritage environments.

Keywords: Heritage Building Information Modeling (HBIM), NoSQL, Graph Database, CIDOC CRM

DOI: 10.18154/RWTH-CONV-254892. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

The use of digital technologies has become central to the documentation and preservation of cultural heritage. Heritage Building Information Modeling (HBIM) offers a structured framework to integrate geometric, historical, and restoration data into unified digital representations that support conservation planning and interdisciplinary collaboration [1], [2].

However, current HBIM implementations often rely on rigid relational databases or proprietary software, limiting their ability to manage heterogeneous, evolving datasets and semantic relationships [3], [4]. This results in poor interoperability, restricted data reuse, and difficulties in long-term access [5], [6].

Ontology-based approaches such as CIDOC CRM offer a solution by enabling structured, semantically rich modeling of heritage knowledge [7], [8]. Yet, technical barriers and missing infrastructure have hindered their broader adoption in HBIM contexts [9].

This paper proposes a hybrid NoSQL system architecture that combines flexible document storage (MongoDB) with semantic graph modeling (Neo4j). The aim is to create a scalable, modular platform for data integration, versioning, and semantic querying—aligned with FAIR principles and cultural heritage ontologies [10], [11]. Rather than replacing existing HBIM platforms, this approach offers a complementary foundation for semantically enriched and interoperable heritage data systems.

2 Related Work

2.1 From Geometric Models to Semantically Enriched HBIM

Historic Building Information Modeling (HBIM) has developed from a tool for geometric reconstruction into a semantically enriched infrastructure for documenting and analyzing cultural heritage. Early methods translated point clouds into parametric models [1], while more recent approaches integrate historical data, material attributes, and restoration processes into structured, queryable models. This shift enables HBIM to function not only as a visual model but as a knowledge system, facilitating reasoning, data reuse, and interoperability [3].

2.2 Structural and Semantic Limitations of Relational Data Models

Relational databases are still widely used in HBIM to store attributes of building elements. However, their rigid schemas hinder integration of diverse data types—such as archival documents, 3D geometries, or annotations—and complicate modeling of temporal or interpretive relationships. Representing changes across restoration phases or aligning heterogeneous metadata often requires workarounds that compromise semantic clarity [5]. Critically, relational systems are not natively compatible with semantic web standards like Resource Description Framework (RDF), limiting their interoperability and long-term usefulness.

2.3 NoSQL Data Models for Flexible HBIM Workflows

NoSQL databases address many of the structural limitations of relational systems. Document-oriented systems like MongoDB support schema-flexible storage of HBIM records—including geometries, pathologies, and intervention logs—within nested data structures. Graph databases such as Neo4j or ArangoDB, in turn, provide native capabilities for modeling historical, material, and actor-based relationships via subject–predicate–object semantics. This aligns closely with ontology-driven frameworks and allows for advanced querying across semantically linked data [9].

2.4 Semantic Enrichment and Ontology Integration in HBIM

The use of ontologies is increasingly central in HBIM to represent knowledge beyond geometric attributes. CIDOC CRM has become a widely accepted standard for modeling events, actors, materials, and sources in a semantically interoperable way. For example, elements like wall fragments (E22 Man-Made Object) can be linked to specific conservation actions (E11 Modification) and responsible agents (E39 Actor), facilitating semantic tracing of interventions across time [7].

2.5 Towards FAIR and Interoperable HBIM Infrastructures

The FAIR principles—Findable, Accessible, Interoperable, and Reusable—have become a guiding framework for managing cultural heritage data. In this context, RDF graphs, persistent URIs, and alignment with shared vocabularies (such as CIDOC CRM) enable HBIM data to be integrated with external repositories and semantic web platforms. While such approaches are gaining traction in the digital humanities, their systematic application within HBIM environments remains limited [8].

2.6 Research Gap in Semantic HBIM Systems

Despite increased awareness of semantic data needs, most HBIM systems continue to rely on proprietary platforms or relational backends, limiting interoperability and adaptability. Hybrid NoSQL architectures—combining flexible document storage with graph-based reasoning—have shown potential in related fields but are rarely operationalized in heritage practice. Coordinated use of MongoDB and Neo4j, aligned through CIDOC CRM, offers a promising path forward, which this research develops into a modular infrastructure [10].

3 System Architecture and Methodology

This section introduces a modular HBIM architecture that combines flexible document storage, semantic graph reasoning, and interoperable metadata standards. It addresses the limits of relational databases and fragmented NoSQL use by enabling structured ingestion and querying of heterogeneous data—from 3D scans to historical records. A prototype was tested using simulated restoration data and point cloud-based models.

3.1 Conceptual Framework

This research proposes a multilayer NoSQL architecture that addresses structural and semantic limitations in conventional HBIM workflows. It combines MongoDB for flexible data storage and Neo4j for semantic querying, enabling scalable integration of heterogeneous data such as IFC models, photogrammetry meshes, annotations, and restoration records.

The system functions as a dynamic information infrastructure—supporting the linkage of restoration data, semantic relationships, and external sources through ontology alignment. Its design is guided by three principles: schema flexibility, semantic expressiveness, and interoperability with established ontologies.

The architecture comprises:

- A document storage layer for ingesting and managing heterogeneous HBIM data
- A semantic graph layer for modeling complex relationships among heritage entities
- An integration layer for coordinating queries and maintaining data consistency between components

This layered approach supports complex queries and enrichment workflows. For example, an element such as `WALL_102` is linked to condition reports, restoration events, materials, and survey data in a coherent and queryable format.

3.2 MongoDB Layer: Flexible HBIM Data Storage

The document storage layer manages core HBIM data using a flexible, schema-less format capable of storing geometry, materials, conditions, and restoration records. Each architectural element is represented as a structured document with fields for identification, spatial context, and intervention history. For example, `WALL_102` includes IFC geometry, STL mesh, materials like limestone, pathologies such as cracking, and associated restoration actions. This structure supports extensibility, versioning, and links to external survey data, adapting to diverse heritage documentation needs.

3.3 Neo4j Layer: Semantic Reasoning and Historical Tracing

The semantic graph layer enables querying relationships such as restoration sequences, material usage across elements, and classification of interventions. Data is aligned with CIDOC CRM, which provides standardized classes for actions (e.g., E11 Modification), building components (E22), and responsible actors (E39), ensuring semantic clarity and interoperability across platforms.

3.4 Interoperability Layer and Query Engine

The architecture enables synchronization between document-based and semantic data by transforming storage-layer content into graph representations while preserving context. This bidirectional structure supports both structured filtering and semantic querying for tasks like restoration planning and documentation analysis.

Although no user interface is included, the modular design supports future integration of dashboards or web tools without altering core logic.

3.5 Standards and Ontologies

Semantic consistency is achieved through CIDOC CRM, which structures data on interventions, materials, spaces, and actors. Metadata is exportable as RDF or JSON-LD (Linked Data), supporting future integration with linked data platforms.

This combination of flexible storage and semantic modeling enhances HBIM workflows, enabling traceability, interoperability, and long-term conservation support.

4 Implementation and Use Case

4.1 Overview

To test the proposed architecture, a prototype was developed to simulate the documentation and analysis of restoration activities associated with historical architectural elements. A representative synthetic data sample was assembled based on common HBIM conditions, including material degradation, restoration actions, and survey data. The goal was to assess the multilayer approach in terms of data structuring, semantic linking, and query execution across both the document and graph-based layers.

4.2 Data Modeling Scope

Several architectural elements were modeled and stored in the document database using a flexible JSON structure. Each record includes geometry (e.g., IFC files, meshes), material data, location metadata, condition annotations, and restoration events with dates, techniques, and responsible actors. A representative example is `WALL_102`, which combines photogrammetric survey data, historical notes (e.g., surface erosion, staining), and restoration metadata. This synthetically constructed case reflects

typical HBIM scenarios and demonstrates the system's capacity for data integration and semantic linkage (see Figure 1).

Figure 1: `WALL_102` Element: Document Structure in MongoDB

In parallel, the graph database was populated with the corresponding semantic entities. These include nodes for elements, restoration events, materials, and techniques, along with relationships such as `RESTORED_ON`, `HAS_MATERIAL`, and `USED_TECHNIQUE`. The resulting structure enables reasoning over intervention histories, material usage, and patterns of deterioration (see Figure 2).

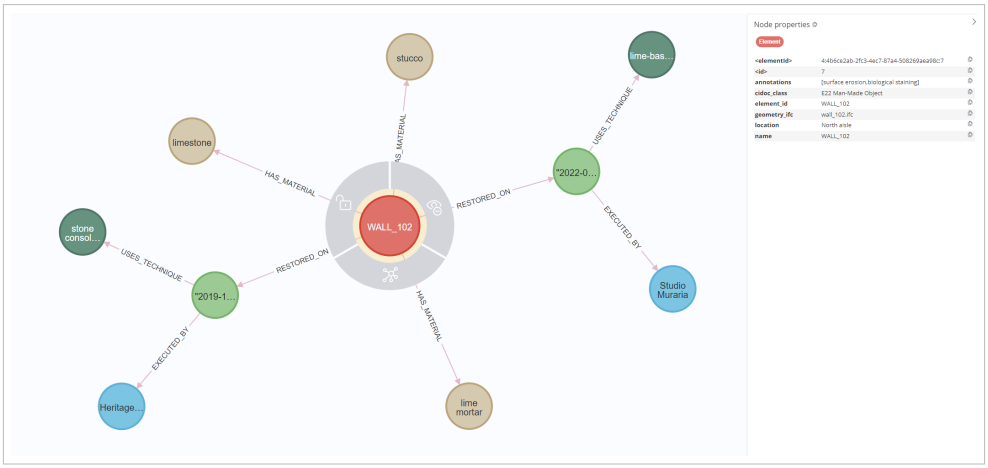


Figure 2: `WALL_102` : Semantic Graph of Restoration Relationships

In addition to individual elements, the system supports the semantic enrichment of full IFC-based digital models. These cases represent a different example from `WALL_102` and are used to demonstrate how restoration metadata, scan data, and geometry can be linked at the model level.

For instance, a file `WALL_test.ifc` includes synthetic metadata connecting the model to STL mesh files, scan techniques (e.g., terrestrial laser scanning), the executing contractor, and a list of linked IFC

elements, see Figure 3 . Each linked element, such as a wall or door, is separately documented with its own restoration history, location, and classification.


```
_id: ObjectId('6849311ab74b62214d174dc6')
file_name : "Wall_test.ifc"
contractor : "BauKultur GmbH"
created_on : "2025-06-11"
▾ linked_elements : Array (2)
    0: "1xLMsTjQr9N88Q3rX5Sot7"
    1: "1xLMsTjQr9N88Q3rX5So9i"
path_ifc : "C:\Users\valde\Documents\05_DenkmalBIM\05_Konferenz\fbi\hbim_prototype…"
path_stl : "C:\Users\valde\Documents\05_DenkmalBIM\05_Konferenz\fbi\hbim_prototype…"
scan_file : "Wall_test.stl"
scan_technique : "Terrestrial Laser Scanning"
source : "academic simulation"
type : "IFC"
```


Figure 3: Digital Model Record for `WALL_test.ifc` in MongoDB

This integration is further represented in the semantic graph layer. The `DigitalModel` node for `WALL_test.ifc` is connected to its STL mesh, scanning process, and the individual IFC components that are part of the model. Each of those components is also linked to restoration events, materials, and responsible actors, enabling complete semantic tracing from the model level down to individual interventions (see Figure 4).

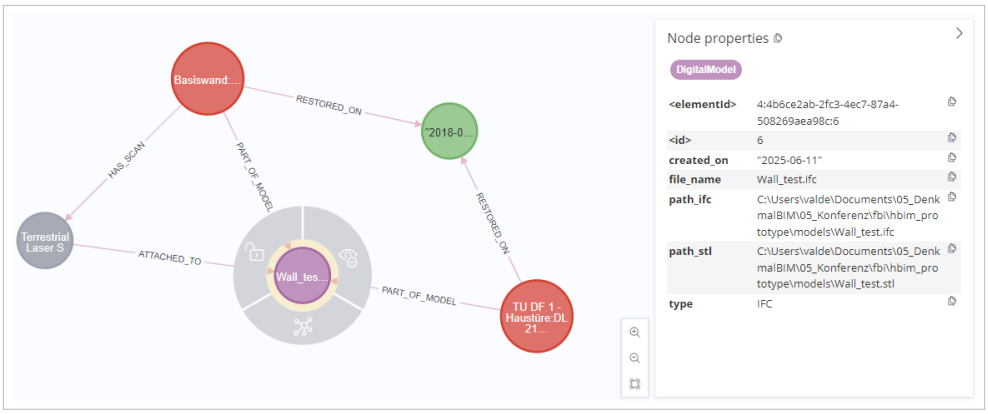


Figure 4: Digital Model Record for `WALL_test.ifc` in MongoDB

This structure allows digital models to act not only as geometric references, but also as information containers that connect technical documentation, historical interventions, and survey data. It supports advanced queries that bridge IFC elements, restoration metadata, and spatial models in an integrated and extensible way.

4.3 Evaluation and Limitations

The prototype validates the feasibility of a hybrid NoSQL architecture for HBIM, enabling structured data management and semantic enrichment of restoration information. It effectively links architectural elements with their historical and material context and offers a scalable foundation for future development. Limitations include the absence of version control, full ontology alignment, and a front-end interface. Coordination between document and graph layers was achieved via scripts rather than automated integration.

Despite these constraints, the system demonstrates clear practical value. Its combination of flexible storage and semantic reasoning exceeds the capabilities of conventional HBIM tools and supports advanced heritage conservation workflows.

5 Conclusions and Future Work

5.1 Conclusions

This research presented and evaluated a hybrid NoSQL architecture designed to support scalable and semantically enriched data management in HBIM. Addressing challenges such as fragmented sources, limited interoperability, and rigid relational models, the system combines document-based flexibility (MongoDB) with semantic graph reasoning (Neo4j) to improve restoration workflows and heritage documentation.

The prototype confirmed the feasibility of managing heterogeneous HBIM data across two layers. Documents support the ingestion of IFC files, point clouds, and restoration records, while the graph layer enables reasoning over events, materials, and actors using CIDOC CRM alignment.

The system addresses key gaps in current HBIM practices, particularly the integration of historical data into structured, queryable models. Though version control and synchronization were not implemented, the architecture anticipates these capabilities and supports future development.

By aligning with open standards and enabling RDF/JSON-LD export, the system offers a reusable and extensible foundation. Compared to conventional HBIM tools, it enhances data interpretation, interoperability, and long-term preservation.

5.2 Future Work

Future development will focus on transforming the prototype into a fully operational platform for real-world heritage documentation. Several key directions include:

- **Semantic and Temporal Enrichment:** Further integration of temporally annotated interventions and event-based modeling will build on the existing use of CIDOC CRM in Neo4j. Future work may include inferencing or historical reasoning features.
- **Interoperable Integration:** Development of standardized middleware will enable consistent connections between components, allowing integration with GIS platforms, IFC viewers, and heritage registries.
- **Linked Data Compatibility:** Connecting to external vocabularies (e.g., Wikidata, Getty AAT) through persistent URIs and SPARQL endpoints would enhance semantic interoperability and linked open data access.
- **User Interface Development:** A visual interface would improve accessibility, allowing heritage professionals to navigate the semantic graph, interact with data, and execute queries more intuitively.
- **Real-World Validation:** Applying the system in partnership with heritage institutions will be essential for assessing scalability, usability, and workflow integration.

In summary, this research establishes a foundation for a modular, semantically enriched HBIM infrastructure. It supports the long-term reuse, interpretation, and interoperability of restoration data—ensuring that interventions are not only documented but meaningfully connected across time and systems.

References

[1] M. Murphy, E. McGovern, and S. Pavia, "Historic building information modelling (hbim) - adding intelligence to laser and image based surveys of european classical architecture", *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 76, no. 4, pp. 89–102, 2013. DOI: 10.1108/02630800910985108

[2] A. Baik and Y. Alshawabkeh, "Harnessing heritage bim for enhanced architectural documentation of ad deir in petra", *Applied Sciences*, vol. 14, no. 11, p. 4562, 2024. DOI: 10.3390/app14114562

[3] F. Giuliani, F. Gaglio, M. Martino, and A. de Falco, "A hbim pipeline for the conservation of large-scale architectural heritage: The city walls of pisa", *Heritage Science*, vol. 12, no. 1, 2024. DOI: 10.1186/s40494-024-01141-4

[4] T. Penjor, S. Banihashemi, A. Hajirasouli, and H. Golzad, "Heritage building information modeling (hbim) for heritage conservation: Framework of challenges, gaps, and existing limitations of hbim", *Digital Applications in Archaeology and Cultural Heritage*, vol. 35, pp. 1–15, 2024. DOI: 10.1016/j.daach.2024.e00366

[5] F. Matrone, E. Colucci, V. de Ruvo, A. Lingua, and A. Spanò, "Hbim in a semantic 3d gis database", *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-2/W11, pp. 857–865, 2019. DOI: 10.5194/isprs-archives-XLII-2-W11-857-2019

[6] M. Schlögl, M. Durco, I. Börner, P. Andorfer, and K. Illmayer, *Connecting the dots: The case of omnipot*, 2020. DOI: 10.1007/978-3-319-63962-8

[7] M. Doerr, K. Schaller, and M. Theodoridou, "Integration of complementary archaeological sources", in *Beyond the Artifact: Digital Interpretation of the Past. Proceedings of CAA2004*, Archaeolingua, 2010, pp. 64–69.

[8] T. Tietz, O. Bruns, H. Fliegl, E. Posthumus, T. Schrade, and H. Sack, "Knowledge graph-basierte forschungsdatenintegration in nfdi4culture", in *DHd2023: Open Humanities, Open Culture: Zenodo*, A. Busch and P. Trilcke, editors, Zenodo, 2023. DOI: 10.5281/zenodo.7748740

[9] F. Michel, "Integrating heterogeneous data sources in the web of data", Ph.D. dissertation, Université Côte d'Azur, 2017.

[10] V. A. Cotella, "From 3d point clouds to hbim: Application of artificial intelligence in cultural heritage", *Automation in Construction*, vol. 152, p. 104 936, 2023. DOI: 10.1016/j.autcon.2023.104936

[11] F. Kräutli and M. Valleriani, "Corpustracer: A cidoc database for tracing knowledge networks", *Digital Scholarship in the Humanities*, vol. 33, no. 2, pp. 336–346, 2018. DOI: 10.1093/llc/fqx047

BIM and Semantic Web: Key Enablers for Nuclear Knowledge Management?

Pascal Mosler[1], Ahmed Han Karapinar[2], Benedikt Malte Hauke Broich[2] and Theo Kastner-Guhl[2]

[1]Institute of Numerical Methods and Informatics in Civil Engineering, Technical University of Darmstadt, Darmstadt, Germany

[2]Department of Computer Science, Technical University of Darmstadt, Darmstadt, Germany

E-mail(s): mosler@iib.tu-darmstadt.de

Abstract: Knowledge management plays a crucial role in nuclear facilities. Due to the diverse business processes to be mapped across individual life cycle phases and the heterogeneous needs of various stakeholders, no uniform platform standard for knowledge management in the field of nuclear technology has yet been established. This paper presents a concept for a software architecture designed to enable digital knowledge management in this context. It is based on an analysis of existing approaches for linking 3D geometries with semantic information and documents, in particular, based on Building Information Modeling, as well as existing information models in nuclear technology. Both disciplines are increasingly relying on semantic web technologies to enable end-to-end interoperability of data across the entire life cycle of a building. Semantic web technologies such as RDF and SPARQL also play a key part in the proposed software architecture. Finally, the paper describes the framework conditions under which the architecture can be further developed into a mature information system. A central feature is the use of extended reality, which enables on-site and collaborative use of the system in multiplayer mode.

Keywords: Knowledge Management, Nuclear Facility, BIM, Semantic Web, Extended Reality

DOI: 10.18154/RWTH-CONV-254893. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

A nuclear facility is *"a facility (including associated buildings and equipment) in which nuclear material is produced, processed, used, handled, stored or disposed of"* [1]. In addition to nuclear power plants (NPPs), these facilities include interim storage facilities, repositories, nuclear reprocessing plants, and research facilities. Assuring the safe and efficient operating phase (operating & maintenance) is crucial for the economic success of such a facility. The workforce of an NPP is large and comprises a wide range of disciplines and qualification levels. A large number of stakeholders are involved in the environment of nuclear facilities, including owners and operators, engineering construction contractors, regulatory bodies, and others [2].

According to the International Atomic Energy Agency (IAEA), knowledge management (KM) is *"an integrated, systematic approach to identifying, managing and sharing an organisation's knowledge [. . .]"* [3]. Knowledge can be divided into explicit knowledge, implicit knowledge, and tacit knowledge. In the case of nuclear facilities, knowledge is continuously accumulated during the individual life cycle phases (planning, construction, operation, decommissioning). Document management is not only an important component of KM but also a crucial part of overall project management [4]. Handling different data formats and attempts to structure data represent a major challenge. Decommissioning is particularly complex for nuclear facilities due to radiation protection and must be carefully planned and carried out. This requires access to the whole operating knowledge of the facility.

2 Technical Basics

2.1 Building Information Modeling

Building Information Modeling (BIM) describes the consistent and continuous use of digital information over the entire life cycle of a built facility—including its planning, construction, and operation [5]. A BIM *model* describes a comprehensive digital image of a building. It contains the three-dimensional geometry of the individual elements of the building as well as non-physical objects such as rooms and floors. The properties and hierarchical dependency of these objects are described semantically [5]. The Industry Foundation Classes (IFC) are a standardized data format or data model for the manufacturer-neutral exchange of digital building models [5]. IFC establishes a strict separation between geometry and semantics. This separation makes it possible to link a semantic object with one (or more than one) geometric representation.

2.2 Semantic Web

The concept of the Semantic Web was introduced by Tim Berners-Lee et al. in 2001 [6]. Based on the World Wide Web (WWW), the aim was to create a machine-searchable system with various machine-readable information and information endpoints. In their conception, they already envisaged the possible linkage between information and physical objects using ontologies [6]. On the one hand, this characteristic makes the Semantic Web interesting in the context of BIM [7]. On the other hand, the Semantic Web plays an important role in knowledge modeling, a subfield of artificial intelligence, where the aim is to achieve machine-readable formalizations [5]. The World Wide Web Consortium (W3C) published the RDF (Resource Description Framework) standard, with which information is formally expressed in triples. A triple is a structural unit that forms a directed, labeled graph, with the edges indicating a named relationship between two resources represented as nodes [8]. Ontologies are largely based on these RDF triples. In analogy to the Structured Query Language (SQL) for relational databases, the W3C has established SPARQL as a query language for graphs.

2.3 (Knowledge-centric) Plant Information Models

In 2009, the American research organization EPRI (Electric Power Research Institute) started to develop a standardized information model, the Plant Information Model (PIM). The background to the development of the PIM was the realization that "silo thinking", i.e. distributed data silos within an organization, leads to unnecessary cost increases due to inefficient collaboration. The PIM was to be used for the construction of new nuclear facilities, enable interoperability, and map the entire plant life cycle. KM has also already been recognized as a use case as well as the advantages of

linking documents and data with a 3D model of the facility based on an object taxonomy. Technically, XML (Extensible Markup Language) was used to enable an object relationship model. In 2014, the development was suspended without the PIM having been completed [9].

In 2020, the IAEA published its concept of a knowledge-centric PIM and defined it as *"a semantically organized set of information describing plant structures, systems and components [...] within a knowledge framework [...]"* [8]. This concept combines the characteristics of a PIM and semantic web technologies. The focus is no longer on mere data, but on contextualized information, and enables more efficient KM internally and in exchange with stakeholders [8].

3 Existing Approaches for Linking Geometry with Semantic Information and Documents

The relationship between geometry and semantics plays a major role in KM. This chapter examines this interplay in the AEC sector and in the nuclear industry.

3.1 AEC Sector

The idea of linking project documents to a model-based information system was already being discussed before BIM was established [10]. An early example is that of Chau et al. [11], in which the Work Breakdown Structure (WBS) was proposed as a link to 3D CAD. In 2019, the early approaches of using the WBS to link objects and information/documents were taken up again by Cha and Lee [12]. In their BIM-based document management system, documents are compiled by the respective owner and organized using a breakdown structure. The documents are stored in a database and their metadata is linked to the BIM [12].

Motawa and Almarshad [13] proposed a knowledge-based BIM for building maintenance in which specific semantic information is entered in the Autodesk Revit authoring software and exported with the IFC schema. A more comprehensive proposal was presented by Opitz et al. [4], who proposed technologies such as RDF and XMP (Extensible Metadata Platform) to link documents and their metadata. In their exemplary application BIMdox, a mapping takes place between the document metadata and the IFC schema. The documents are stored in a database and searched using SQL. In addition, they categorized basic options for creating a link between documents and objects in a project management environment: a) directly in the building model (using IFC), b) external link model, c) as part of an ontological model specification, and d) within the document meta-model.

Quattrini et al. [14] use the example of HBIM (Heritage BIM) to present a Semantic Web-related linking approach in which IFC data is stored in an RDF triplestore. In the Revit authoring software, the user can semantically enrich a BIM via IFC properties and export it in the IFC schema. Using an IFC-to-RDF tool, the information can be converted into triples and entered into a triplestore. Converters such as IFC to RDF play a decisive role in such approaches. For example, Bonduel et al. [15] developed an IFCtoLBD converter based on ifcOWL and the IFC-to-RDF converter, which focuses on linked (building) data graphs. This converter is also used by Jäkel et al. [16] to convert IFC models into RDF graphs. They go on to use Information Container for Linked Document Delivery (ICDD) for validation.

3.2 Nuclear Industry

In 2021, the French energy supplier EDF (Électricité de France) launched the plant life cycle management (PLiM) project. The aim is to improve information management—especially for new reactors. Overall, the central tasks of KM in nuclear facilities are to be addressed, including the internal exchange of knowledge, the exchange with external stakeholders, and the transfer of knowledge to future generations. Technologically, a kind of digital twin shall be created in which data can be consistently generated, validated, stored, and retrieved [8]. Akinyemi et al. [17] developed a data integration framework for offshore decommissioning waste management in the oil and gas domain. There are numerous cross-references between the oil and gas domain and the nuclear industry, which result from similarities in the complexity of the plants and the same safety culture. They analyzed the architecture of existing research approaches in the field of waste management for building assets and identified in the work of Bilal et al. [18] a very advanced approach with the use of Neo4j for data storage and ifcOWL for knowledge representation. In their own proposal, the documents are stored in a relational database, while the meta information is stored in RDF format in a triplestore. This information can be searched using SPARQL. To input the data, they define a spreadsheet template.

Pallas is a reactor for the production of medical isotopes that is currently under construction. Extensive efforts are being made here to test semantic interoperability in practice. The data exchange processes within a Common Data Environment (CDE) should be carried out following the standards RDF(S) and SPARQL. Breakdown structures and the IFC model of the plant form the backbone of the CDE [19].

4 Analysis

This chapter highlights the differences and similarities between the approaches to linking geometry with semantic information and documents in the AEC sector and the nuclear industry.

4.1 Life Cycle Phases and Challenges in the Knowledge Transfer

It is noticeable that KM is closely related to data and information management in practical implementation. Proximity to BIM has only recently been systematically established in nuclear technology through research projects such as Pallas [19]. Semantic web technologies are becoming increasingly important in this context. The IAEA also makes an explicit recommendation for their use in KM and data exchange [20]. In the AEC sector, the life cycle phases are weighted differently than in nuclear facilities. The most striking difference to the phase definition of the IAEA [8] is that decommissioning hardly plays a role in the AEC sector. For nuclear facilities, on the other hand, decommissioning is a decades-long, labor-intensive, and costly process with high strategic relevance. The experiences in the construction industry and in the nuclear industry are comparable in terms of the build-up of knowledge and, in particular, its loss in transition phases.

4.2 BIM and Spatial Anchoring of Knowledge and Documents

BIM is still considered a relatively “young” methodology in nuclear engineering. Here, KM methods are not primarily based on a BIM model, but on considerations of traditional document management. Accordingly, little attention is paid to the interdependencies of geometry and semantics in BIM. Instead, BIM is generally understood as “geometry”. Is BIM-based KM therefore a suitable approach for nuclear facilities? A closer look shows that there are considerable implementation challenges, particularly in the operational phase. Current research and concepts of digital twins are increasingly linking BIM

to the operation phase. Cha and Lee [12] emphasize that information management is particularly important in this phase, but that BIM in its current implementation cannot be the sole anchor point for information management in this phase. Linking documents and knowledge with BIM elements ultimately results in the spatial allocation of this information to physical objects. Mandatory component identification systems are already used in NPPs, which serve as a reference for the analog storage of documents. Particularly during the operating phase, with regular maintenance and revisions to be carried out, the identification system enables the unique allocation of physical components —not only for analog document management purposes but perspectively also for digital KM.

4.3 Ontologies and Data Structures for Integrating Explicit and Implicit Knowledge

In the research and practical approaches considered in this work, there is little differentiation between documents, data, and knowledge. In classical KM, however, clear boundaries are drawn. The mere attribution of a BIM model in authoring software, as in Motawa and Almarshad [13] and Deshpande et al. [21] may or may not be equivalent to the input of knowledge [22]. The basic principle of the Semantic Web of mapping and exchanging knowledge at a granular level can also be used for mere information or data. Implicit knowledge plays a prominent role in the economic success of an organization [22]. However, current documentation requirements in both the construction and nuclear industries refer to explicit knowledge. Sobhkhiz and El-Diraby [23] point out that documented implicit knowledge often exists in unstructured data form, is difficult to integrate into existing structures with predominantly explicit knowledge and therefore remains predominantly in the heads of the knowledge carriers. Hu et al. [22] point out that even with highly developed knowledge organization systems such as ontologies and other technologies such as knowledge graphs, the hurdles to mapping experiential/implicit knowledge are high. For Sobhkhiz and El-Diraby [23] however, ontologies alone are no guarantee that unstructured data can be better linked. They point to the maintenance and further development effort for such structures, especially if relations between entities change in the future.

5 Concept

In the following, a software architecture for a fully digital KM platform for nuclear facilities is outlined. In earlier research [24], extended reality (XR) was discussed as a visualization technology in nuclear technology and was seen as a useful solution for enabling direct KM on-site. With the help of XR, spatially linked knowledge can be viewed on-site (AR), but a location-independent experience (VR) is also possible. The implementation in software is possible with the Unity Game Engine. A suitable headset that is both AR- and VR-capable is the Meta Quest 3. To view knowledge—meaning documents and granular knowledge elements linked to physical components reflected in a BIM model—together, a multiplayer mode must be established. This also enables verbal exchange, which is of great importance in analog KM settings. Meta provides corresponding functions for the Meta Quest 3 with the Colocation SDK, which are supplemented by the Photon Unity multiplayer network.

Given these constraints, how could data handling be organized? The lowest common denominator of the considered Semantic Web approaches is that they use RDF as their data model. The concept by Akinyemi et al. [17] considered in Section 3.2 proposes a sensible connection and storage of semantic data with regard to information management. This includes the outsourcing of binary files, a triplestore

for storing metadata, and a document meta-model for formalizing knowledge. Such a model with defined attributes can be used as a spreadsheet template for entering existing data, but also later as an “interface” for the user to insert implicit knowledge elements within the application.

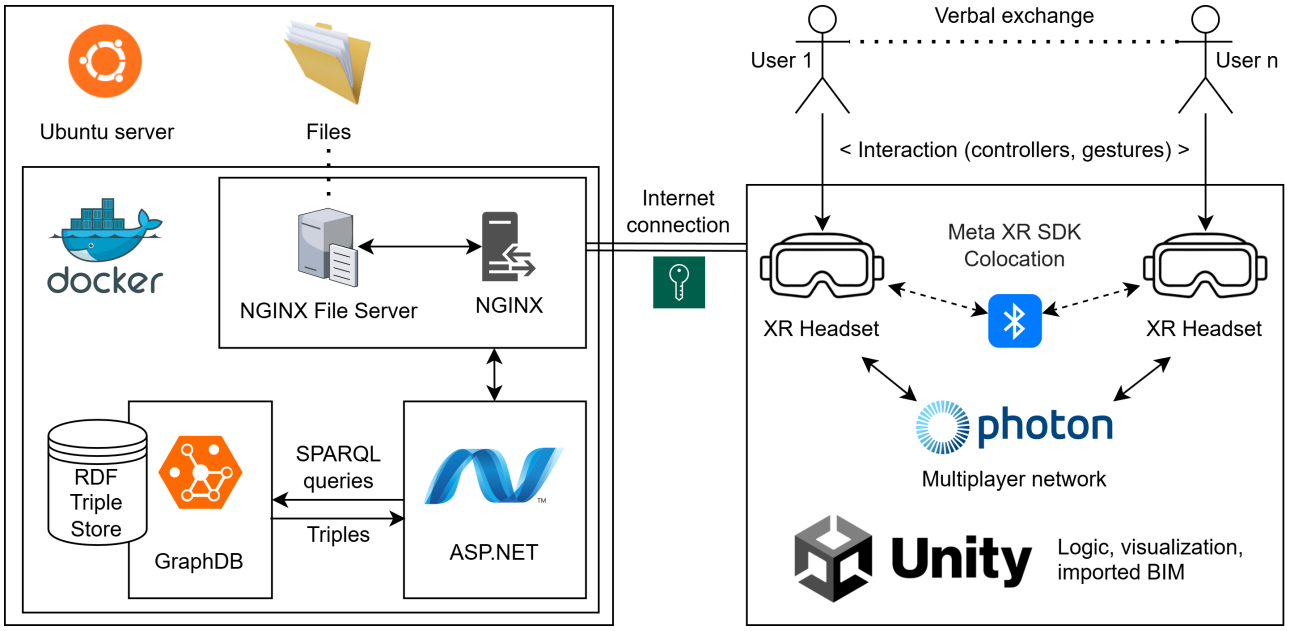


Figure 1: Proposed software architecture for an XR-based KM platform for nuclear facilities

In the software architecture presented here, an RDF triplestore is connected to Unity. The dotNetRDF framework is used to query and manipulate RDF data. A BIM model can either be deposited in Unity or linked to it. Depending on the method chosen, the component attributes are retained in the IFC schema and can be read out in Unity. The management of knowledge elements and binary data as well as their linking to geometries is carried out on an Ubuntu server using the document meta-model. The server uses Docker as an abstraction and security layer. There are three Docker containers: NGINX (file server and reverse proxy), ASP.NET (web framework for creating a REST API), and GraphDB (RDF triplestore). Figure 1 shows the software architecture developed. RDF is used in numerous ontologies and vocabularies, is flexible and expandable. Therefore, due to the RDF data structure, ontologies can be added as required without having to change the basic model.

6 Discussion and Conclusion

In this contribution, an XR-based KM platform for use in nuclear facilities was proposed incorporating established principles from BIM and the Semantic Web. In future research, this software architecture is to be further developed into a prototypical information system and tested in a suitable environment. The overall goal is to establish a higher-level approach that can be adapted and refined based on existing geometric and semantic data, depending on the specific use case. Hu et al. [22] see potential in the higher-level field of knowledge science through technologies such as IoT, AI, cloud computing, and blockchain. The platform presented here can generally be expanded using these technologies. In particular, AI functionalities such as RAG (Retrieval-Augmented Generation) could usefully supplement the classic querying of information from the graph database and binary data.

References

[1] International Atomic Energy Agency, *IAEA Nuclear Safety and Security Glossary: Terminology Used in Nuclear Safety, Nuclear Security, Radiation Protection and Emergency Preparedness and Response* (Non-Serial Publication), 1st ed. Vienna: International Atomic Energy Agency, 2022.

[2] International Atomic Energy Agency, *Initiating nuclear power programmes: Responsibilities and capabilities of owners and operators: Iaea nuclear energy series no. ng-t-3.1 (rev. 1)*, 2020.

[3] International Atomic Energy Agency, *Guide to knowledge management strategies and approaches in nuclear energy organizations and facilities: Iaea nuclear energy series no. ng-g-6.1*, Vienna, 2023.

[4] F. Opitz, R. Windisch, and R. J. Scherer, "Integration of document- and model-based building information for project management support", *Procedia Engineering*, vol. 85, pp. 403–411, 2014. DOI: 10.1016/j.proeng.2014.10.566

[5] A. Borrmann, M. König, C. Koch, and J. Beetz, editors, *Building information modeling: Technologische Grundlagen und industrielle Praxis* (VDI-Buch), 2., aktualisierte Auflage. Wiesbaden and Heidelberg: Springer Vieweg, 2021.

[6] T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic web: A new form of web content that is meaningful to computers will unleash a revolution of new possibilities", *Scientific American*, vol. 284, no. 5, pp. 34–43, 2001. DOI: 10.1038/scientificamerican0501-34

[7] P. Hagedorn et al., "Bim-enabled infrastructure asset management using information containers and semantic web", *Journal of Computing in Civil Engineering*, vol. 37, no. 1, 2023. DOI: 10.1061/(ASCE)CP.1943-5487.0001051

[8] International Atomic Energy Agency, *Application of plant information models to manage design knowledge through the nuclear power plant life cycle* (IAEA TECDOC Series). Vienna: IAEA, 2020, vol. no. 1919.

[9] EPRI, "Advanced nuclear technology: New nuclear power plant information turnover guide (revision 1)",

[10] C. H. Caldas, L. Soibelman, and L. Gasser, "Methodology for the integration of project documents in model-based information systems", *Journal of Computing in Civil Engineering*, vol. 19, no. 1, pp. 25–33, 2005. DOI: 10.1061/(ASCE)0887-3801(2005)19:1(25)

[11] K. W. Chau, M. Anson, and J. P. Zhang, "4d dynamic based management and visualization software: 1. development", *Automation in Construction*, vol. 14, no. 4, pp. 512–524, 2005.

[12] H. Cha and D. Lee, "Framework based on building information modelling for information management by linking construction documents to design objects", *Journal of Asian Architecture and Building Engineering*, vol. 17, no. 2, pp. 329–336, 2018. DOI: 10.3130/jaabe.17.329

[13] I. Motawa and A. Almarshad, "A knowledge-based bim system for building maintenance", *Automation in Construction*, vol. 29, pp. 173–182, 2013. DOI: 10.1016/j.autcon.2012.09.008

[14] R. Quattrini, R. Pierdicca, and C. Morbidoni, “Knowledge-based data enrichment for hbim: Exploring high-quality models using the semantic-web”, *Journal of Cultural Heritage*, vol. 28, pp. 129–139, 2017. DOI: 10.1016/j.culher.2017.05.004

[15] M. Bonduel, J. Oraskari, P. Pauwels, M. Vergauwen, and R. Klein, “The ifc to linked building data converter : Current status”, 2018.

[16] J.-I. Jäkel, E. Heinlein, H. Morgenstern, H. Kim, and K. Klemt-Albert, “System- and data-integrated linking of digital 3d models of existing bridge structures with knowledge graphs of non-destructive diagnostic methods”, *Journal of Information Technology in Construction*, vol. 30, pp. 603–630, 2025. DOI: 10.36680/j.itcon.2025.025

[17] A. G. Akinyemi, M. Sun, and A. Gray, “Data integration for offshore decommissioning waste management”, *Automation in Construction*, vol. 109, p. 103 010, 2020. DOI: 10.1016/j.autcon.2019.103010

[18] M. Bilal et al., “Big data architecture for construction waste analytics (cwa): A conceptual framework”, *Journal of Building Engineering*, vol. 6, pp. 144–156, 2016. DOI: 10.1016/j.jobe.2016.03.002

[19] L. van Ruijven, *The pallas project - a practical experience of semantic interoperability: Complexity and capabilities of skos, rdf and owl/shacl*, 2023.

[20] International Atomic Energy Agency, *Exploring Semantic Technologies and Their Application to Nuclear Knowledge Management* (IAEA Nuclear Energy Series). Vienna: IAEA, 2021.

[21] A. Deshpande, S. Azhar, and S. Amireddy, “A framework for a bim-based knowledge management system”, *Procedia Engineering*, vol. 85, pp. 113–122, 2014. DOI: 10.1016/j.proeng.2014.10.535

[22] Z.-Z. Hu, S. Leng, J.-R. Lin, S.-W. Li, and Y.-Q. Xiao, “Knowledge extraction and discovery based on bim: A critical review and future directions”, *Archives of Computational Methods in Engineering*, vol. 29, no. 1, pp. 335–356, 2022. DOI: 10.1007/s11831-021-09576-9

[23] S. Sobhkhiz and T. El-Diraby, “Dynamic integration of unstructured data with bim using a no-model approach based on machine learning and concept networks”, *Automation in Construction*, vol. 150, p. 104 859, 2023. DOI: 10.1016/j.autcon.2023.104859

[24] P. Mosler and U. Rüppel, “Designing a knowledge-centric plant information model by integrating tacit knowledge”, *31st International Conference on Nuclear Engineering*, 2024. DOI: 10.1115/ICONE31-135585

Verbesserung des Informationsflusses in der Infrastrukturplanung: Ein systemtheoretischer Ansatz

Bastian Wons[1] and Martin Wogan[1,2]
[1]B/M Consult GmbH, Theodor-Heuss-Straße 9a, 38122 Braunschweig, Germany
[2]Institut für Bauinformatik, Technische Universität Dresden, Germany
E-mail(s): bastian.wons@bmconsult.de, martin.wogan@bmconsult.de

Abstract: Die Planung von Infrastrukturprojekten erfordert die Berücksichtigung vielfältiger Anforderungen, die von Stakeholdern über den gesamten Lebenszyklus hinweg gestellt werden. Diese Vielschichtigkeit führt zu einem enormen Informationsaufkommen, welches aufgrund von Kommunikationsverlusten und einer unüberschaubaren Datenmenge schwer zu filtern ist. Planer und Bauleiter stehen daher vor der Herausforderung, entscheidungsrelevante Informationen effizient zu identifizieren und zu nutzen. Besonders in innerstädtischen Infrastrukturprojekten zeigt sich ein Defizit in der Verbreitung digitaler Planungsmethoden, was die Bewältigung der Informationsfülle zusätzlich erschwert. Um diesen Herausforderungen zu begegnen, wird in dieser Arbeit eine Methode vorgestellt, die den Informationsfluss innerhalb der Planungs- und Ausführungsphase verbessert. Basierend auf der Systemtheorie wird das Bauprojekt als ganzheitliches System betrachtet, aus dem die Bedürfnisse der Stakeholder mittels profilbezogenen LOIN (pLOIN) abgeleitet und in konkrete Anforderungen an das Infrastrukturprojekt überführt werden. Diese Anforderungen dienen als Grundlage für ein planungsphasenabhängiges LOIN (ppaLOIN), das die relevanten Informationen gezielt bereitstellt und somit die Informationsflut reduziert. Die vorgeschlagene Methode zielt darauf ab, die Informationsbereitstellung in Infrastrukturprojekten zu strukturieren, Entscheidungsprozesse zu erleichtern und somit eine effizientere sowie bedarfsgerechtere Planung und Umsetzung zu ermöglichen.

Keywords: Lebenszyklusprozess, Level of Information Need, Informationsmanagement, BIM, Infrastrukturplanung

DOI: 10.18154/RWTH-CONV-254894. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Einleitung

Infrastrukturprojekten durchlaufen drei zentrale Phasen: Planungs-, Ausführungs- und Nutzungsphase. In jeder dieser Phasen werden Informationen erzeugt, gespeichert, ausgetauscht und überarbeitet, um das Handeln im Projektablauf festzulegen. Entscheidungsprozesse innerhalb dieser drei Phasen basieren auf der Abwägung verschiedener Handlungsoptionen, deren potenzieller Konsequenzen sowie relevanter Zielgrößen und Randbedingungen [1]. In der Praxis zeigt sich, dass Entscheidungen,

respektive die damit verbundenen Planungsleistungen, häufig überarbeitet werden müssen, da neue Informationen erzeugt werden, Informationen aufgrund von Informationsverlusten nicht in der Planung des Projekts berücksichtigt wurden (zeitlicher Versatz) oder sich innerhalb des Planungsprozesses Anforderungen ändern. Dies kann zur Verbesserung des Projekts beitragen, führt jedoch häufig zu Verzögerungen im Planungsfortschritt, sowie zu steigenden Kosten durch vermeidbare Iterationsschritte betreffend der Planungsinhalte [2, 3]. Eine schnelle Entscheidung, ob in der Planung oder Ausführung, kann somit nicht getroffen werden, da den Entscheidern nicht unmittelbar alle relevanten Informationen zur Verfügung stehen. Diese Untersuchung zeigt auf, wie eine mögliche Verknüpfung von Informationen innerhalb von Infrastrukturplanung zum Informationsaustausch nach dem Vorbild der BIM-Methode gestaltet werden kann. Die Entwicklung der Systemanforderungen basiert auf einem strukturierten Requirement-Engineering-Prozess gemäß ISO 29148, der bislang nach aktuellem Stand der Literatur nicht spezifisch für Infrastrukturprojekte angewendet wurde. Dies gilt ebenso für die Ableitung und Entwicklung einer Systemarchitektur in Anlehnung an die Systemkonzepte nach Ropohl und Köchendörfer [4, 5], die zur Abbildung des Informationsflusses innerhalb einer Infrastrukturplanung entsprechend im Rahmen des Verständnisses angewendet wurden. Daher ist die Methodik an dieser Stelle an die gegebene Literatur angelehnt und entsprechend für die Ziele des abzubildenden Produktsystems einer Infrastrukturplanung interpretiert worden.

2 Stand der Forschung

Innerstädtische Infrastrukturprojekte werden überwiegend von kleinen und mittelständischen Ingenieurbüros bearbeitet. Primär wird mit Computer Aided Design (CAD) -Software und speziellen Trassierungsprogrammen für den Gleis- und Straßenbau geplant. Der Fokus dieser Anwendungen liegt vorrangig auf der Ableitung von 2D-Plänen, während modellbasierte Planungsmethoden (Building Information Modelling (BIM)) nur begrenzt zum Einsatz kommen [6]. Ein zentrales Konzept innerhalb der BIM-Methodik ist das „Level of Information Need“ (LOIN) [7]. Dieses Konzept verfolgt das Ziel, Informationen bedarfsgerecht bereitzustellen und gleichzeitig eine übermäßige Datenflut zu vermeiden [8]. Hierbei werden nur die tatsächlich für den jeweiligen Prozess erforderlichen Informationen durch die jeweiligen Planer zur Verfügung gestellt, was sicherstellt, dass Planungs- und Entscheidungsprozesse effizient gestaltet werden können und keine “überflüssigen” Daten erzeugt werden. Die Verfügbarkeit präziser und bedarfsgerechter Informationen spielt eine zentrale Rolle für die Qualität und das Verständnis von Entscheidungsprozessen während der Planung und Ausführung von Infrastrukturprojekten und bedingt somit, dass der Projektlebenszyklus ein Informations-Lebenszyklus-Management (ILM) erfordert. [9, 10, 11, 12]

2.1 Informations-Lebenszyklus-Management

Information Requirements (Informationsanforderungen, IR) nach DIN EN ISO 19650-1 sind die Grundlage für den durchgängigen Austausch von Informationen in einer Organisation beziehungsweise für die Abwicklung von Projekten. Um Organisations- und Projektziele zu erreichen, benötigt der Informationsbesteller (AG oder Planer untereinander) Informationen. Die Aufgabe des Informationsbestellers ist es, die IR gegenüber dem Informationsbereitsteller so zu spezifizieren, wie sie für das Projekt benötigt werden. Dies geschieht innerhalb verschiedener IR, in denen festgelegt wird, »was, wann, wie und für wen« an Informationen erstellt werden muss [13]. Durch iterative Feedbackprozesse und

dynamische Anpassungen des Projektleiters oder Informationsmanagers an die jeweiligen Anforderungen der Projektphasen kann sichergestellt werden, dass nur die zur Projektabwicklung benötigten Informationen ausgetauscht werden. (vgl. [2, 13, 14]).

2.2 Use Case Management

Die wiederkehrende Abfrage von IR kann als Use Case beschrieben werden. Um die Nutzbarkeit der Use Cases zu gewährleisten, müssen die Anforderungen an den Informationsaustausch über den gesamten Lebenszyklus definiert werden. Aufgrund der normierten Struktur der Informationsanforderungen können die Use Cases innerhalb eines Use Case Management Service Systems für die globale Bauindustrie zugänglich gemacht werden [13]. Ein Medium dafür bietet beispielsweise das BIM-Portal des Bundes oder BuildingSMART International [15]. Somit ist es möglich, dass der Informationsbesteller eine Auswahl über die für das Projekt relevanten Use Cases trifft und diese in den IR zu referenzieren. Die so gestellten IR können nach Vergabe des Projekts und Bestellung der Leistung projektspezifisch präzisiert und ergänzt werden vgl. [16].

3 Methode

Im Rahmen dieser Arbeit wurden Entscheidungsprozesse in der Planungs- und Ausführungsphase von Infrastrukturprojekten analysiert und mittels profilbezogene LOIN (pLOIN) eine Methode entwickelt, die es ermöglicht, die Entscheidungsfindung zu beschleunigen. pLOIN kann als eine Art Filter betrachtet werden, der nicht nur den Umfang der weitergegebenen Informationen auf die Relevanten begrenzt, sondern gleichzeitig eine auf den Stakeholder bezogen präzise und zielgerichtete Bereitstellung der Informationen zur Entscheidungsfindung sicherstellt. Dabei ist es essentiell, zwischen Entscheidungen zu differenzieren, die als integraler Bestandteil der Planungsphasen und der damit verbundenen iterativen Prozesse notwendig sind, und solchen, die vermeidbar gewesen wären.

3.1 Entscheidungsprozess

Wichtig für die Entwicklung des pLOIN ist das Verständnis über Entscheidungen, die im Rahmen eines Projektes getroffen werden. Entscheidungen sind erforderlich, wenn es „Optionen" (synonym: Alternativen) gibt. [1] In einer Entscheidungssituation liegen mindestens zwei Optionen vor. Die hier thematisierten Optionen können entweder Handlungen oder Dinge sein, die für die Entscheidung zur Auswahl stehen. Optionen stellen die wesentliche Komponente von Entscheidungsproblemen dar. Auf eine ausführliche Beschreibung der zu berücksichtigenden Kategorien im Entscheidungsprozess wird hier aus Platzgründen verzichtet und auf die Werke von Pfister [1] verwiesen. Aus den Definitionen nach Pfister [1] lassen sich Entscheidungen innerhalb der Planungs- und Ausführungsphase von innerstädtischen Infrastrukturprojekten als reflektierte Entscheidungen einordnen, die insbesondere in der Ausführungsphase eine schnelle Zeitdauer erforderlich machen ("Stress auf der Baustelle"). Entscheidend dafür ist, dass zwischen eindeutig definierten Optionen entschieden werden muss, die für jede Situation immer wieder neu entwickelt werden müssen. Diese Art von Entscheidungen ist mit einem hohen kognitiven Aufwand zur Entscheidungsfindung verbunden und kann daher längere Zeit dauern.

3.2 profilbezogene LOIN (pLOIN)

Die Bedürfnisse der Stakeholder im Lebenszyklusprozess [9] bilden die Grundlage für die Ableitung eines profilbezogenen LOIN (pLOIN). pLOIN stellt ergänzend zum LOIN eine direkte Verknüpfung zu

den jeweiligen Stakeholdern her vgl. [17]. Informationsanforderungen, die z.B. durch den AG innerhalb der IR definiert werden, werden einzelnen Stakeholder oder Stakeholdergruppen zugewiesen und somit zusätzlich strukturiert und differenziert. Somit erfolgt eine zweistufige Aufteilung: Einerseits werden die für den IR relevanten Informationsanforderungen definiert, andererseits wird eine gezielte Zuordnung zu den jeweiligen Stakeholdern vorgenommen. Beispielsweise lassen sich aus dem Informationsmodell genau jene Informationen extrahieren, die in einer Visualisierung (Anwendungsfall „040 Visualisierung“ nach BIM Deutschland (2025)) verdeutlichen, an welchen Stellen (Objekte, Subsysteme, Bauteile) die Anliegen der Bürger in der bisherigen Infrastrukturplanung berücksichtigt wurden.

3.2.1 Beispiel 1: Planungsphase

Das reale Beispielprojekte aus der Planung ist ein innerstädtisches Vorhaben in unmittelbarer Nähe eines komplexen Verkehrsknotenpunktes. Der betrachtete Planungsabschnitt umfasst den Südarm eines Knotenpunktes und stellt eine besondere städtebauliche Herausforderung dar, da sich entlang der westlichen Seite eine bestehende Baumreihe befindet. Gleichzeitig ist auf der östlichen Seite eine neue Bebauung geplant, wodurch der verfügbare Planungsraum begrenzt wird. Die für den Neuentwurf erarbeiteten Informationen wurden gezielt nach relevanter Projektinformationen durchsucht, die zu einer Planungsentscheidung geführt haben. Von besondere Bedeutung waren dabei Vermessungsdaten, digitale Geländemodelle, Schall- und Verkehrsgutachten, Definitionen der Planungsziele sowie vereinzelte spezifischen Anforderungen und Vorgaben des Auftraggebers. Insbesondere über die Zuweisung der vereinzelte spezifischen Anforderungen zu den Profilen von z.B. Stadtgrün, Radbeauftragten, Behindertenvertretung und Versorgern konnten für die Entscheidungen relevanten Information schnell gefunden, priorisiert und für die Entscheidung aufbereitet werden. Iterative Überarbeitungen von Plänen konnten so verhindert werden. Diverse Änderungen der Randbedingung zum Erhalt der Baumreihe, führt schlussendlich dennoch zu einem Planungsstopp und einer Wiederholung der Vorplanung.

3.2.2 Beispiel 2: Bauausführung

Das Projekt befand sich in der Ausführungsphase, was bedeutet, dass bereits alle planerischen Entscheidungen getroffen wurden. Bei auftretenden Ausführungsproblemen war der Projektingenieur für die Entwicklung und Implementierung geeigneter Lösungen verantwortlich, um den Baufortschritt sicherzustellen. Das Projekt umfasst den Neubau von Infrastrukturelementen des öffentlichen Personennahverkehrs (ÖPNV) sowie die Ertüchtigung und Anpassung angrenzender Verkehrsinfrastruktur innerhalb eines bereits bestehenden Straßenraums. Die Anbindung der Planung an den Bestand, stelle hierbei einen Herausforderung dar. Während der Bauausführung musst festgestellt werden, dass ein Bestandsbord im Bereich einer Querung nicht korrekt aufgemessen worden war. Dies führte zu einer Höhendifferenz zwischen dem neu zu errichtenden barrierefreien Querungssteinsystem und dem bestehenden Haltestellen-Bord. Aufgrund dieser Abweichung konnte der Bau nicht wie geplant fortgesetzt werden, was eine Störung des Bauablaufs zur Folge hatte. Mit der Meldung des Problems setzte der Entscheidungsprozess für den Projektingenieur ein. Auf Basis der übermittelten Informationen sollte zunächst eine Einschätzung der Situation erfolgen. Eine Ferndiagnose erwies sich als unzureichend, da keine Bilder der Situation übermittelt wurden und die sprachliche Beschreibung nicht hinreichend war, um die Situation weiter bescheiden zu können. Dies stellt den

Beginn des Informationsbeschaffungsprozesses dar. Nach der Analyse der baulichen Gegebenheiten mussten sämtliche relevanten Planungsunterlagen überprüft werden. Dabei zeigte sich, dass eine inkonsequente und unstrukturierte Informationsverwaltung erhebliche Verzögerungen verursachte. Es fehlt eine zentrale, geordnete Ablage aller entscheidungsrelevanten Dokumente. Es mussten alle Projektunterlagen manuell durchsucht oder Rückfragen an die verantwortlichen Planenden gestellt werden. Dies verlängert nicht nur die Phase der Informationsbeschaffung, sondern verzögert auch die eigentliche Phase der Entwicklung von Lösungsvarianten und somit auch die Entscheidungsfindung. Nachdem Lösungsvarianten erarbeitet und dokumentiert worden waren, beispielsweise in Form von Plänen oder Skizzen, mussten diese durch die Bauleitung vorgestellt und anschließend dem AG zur Entscheidung vorgelegt werden. Der Umfang der Dokumentation richtet sich dabei nach der Tragweite der Anpassung. Während für kleinere Änderungen oft einfache Skizzen mit den wesentlichen Parametern genügen, sind bei komplexeren Maßnahmen detaillierte Pläne erforderlich. Entscheidend ist, dass der Lösungsprozess möglichst zügig erfolgt, um den Baufortschritt nicht unnötig zu beeinträchtigen. Wird ein detaillierter Kostenvoranschlag benötigt oder sind potenzielle Auswirkungen auf andere Projektbereiche unklar, kann sich der Entscheidungsprozess auf der Auftraggeberseite entsprechend verlängern. Nach der Freigabe einer Lösungsoption wurden die angepassten Ausführungsunterlagen an die Baufirma übermittelt, sodass der Bau fortgesetzt werden konnte.

4 Ergebnisse

Die Beispiele 1 und 2 veranschaulichen die Bedeutung von struktorierten Projektinformationen. Durch den pLOIN-Ansatz hätten die Entscheidungen aus der Planung mit den entsprechenden Anforderungen dokumentiert und auf den selben Grundlagen neue Entscheidungen erarbeitet werden können, ohne den Prozess der Informationsbeschaffung neu durchführen zu müssen. Die Bauverzögerung hätte aus Sicht der Autoren deutlich verkürzt werden können. Entscheidungen basieren auf wiederkehrenden Anforderungsmustern, die je Stakeholder unterschiedlich sind. Es können mit dem pLOIN-Ansatz Entscheidungsprozesse modelliert werden, die eine Entscheidungsfindung und somit Lösung von alltäglichen Fragestellungen unterstützen.

5 Diskussion

Es muss weiter untersucht werden, wie die hier gewählte Herangehensweise für die Praxis implementiert werden kann. Die Softwaretechnik-Implementierung in einem DMS steht noch aus. Dies könnte in Anlehnung an das Konzept der Mobilitätsbewertungsplattform (MBPF) [17] erfolgen. Besonders bedeutsam für die Implementierung ist die Verknüpfung von Anforderungsattributen mit den Elementattributen des Systems, da diese direkt die Eigenschaften des Modells sowie der einzelnen Bauteile widerspiegeln. Auf welcher Ebene dies umgesetzt werden kann, muss in weiteren Arbeiten erforscht werden. Eine weitere Herausforderung ist, dass die derzeit etablierten Arbeitsweisen und genutzten Softwarelösungen nicht über die notwendigen Funktionalitäten (manuelle Erweiterbarkeit) oder das erforderliche Fachwissen (offene Schnittstellen zur Anbindung externen Datenbanksysteme) für eine präzise Definition von pLOIN verfügen. Die Zuordnung von Informationen über pLOIN ermöglicht es, die Bedürfnisse von Stakeholder zielbezogen zu befriedigen. Dies ist in der Individualität des Informationsbedarfs jedes Einzelnen begründet, weshalb weitere Beschränkungen, beispielsweise schon

innerhalb des AWF oder durch die Kategorisierung der visualisierten Informationen, benötigt werden. Gleichzeitig bleibt zu berücksichtigen, dass aufgrund der individuellen Herausforderungen, mit denen Bauleiter konfrontiert sind, die Definition von Systemanforderungen keinen unmittelbaren Einfluss auf das LOIN und deren Arbeitsprozesse hat. Am Ende muss weiterhin der Mensch entscheiden. Dabei spielen Erfahrung und das sprichwörtliche Bauchgefühl ebenso eine Rolle, wie die technischen Unterstützungen, die heutzutage zur Verfügung stehen. Nur der Planer bzw. Bauleiter selbst kann festlegen, welche Informationen er zur Lösung einer Fragestellung benötigt. Projektmanagement-Plattformen können hier nur ein weiteres Hilfsmittel sein, um die steigende Komplexität von Bauprojekten im Griff zu behalten. Darüber hinaus stellt sich die Frage, ob pLOIN grundsätzlich für jede Leistungsphase manuell formuliert werden kann oder ob ein selbstlernender Mechanismus zwischengeschaltet werden muss, um den hohen Aufwand einer manuellen Implementierung zu vermeiden. Sollte dies nicht ohne Weiteres möglich sein, wären weitere Untersuchungen nötig, um eine Methode zu entwickeln, die es ermöglicht, Stakeholder-spezifische Informationen so aufzubereiten, dass der individuelle Informationsbedarf gesteuert werden kann.

6 Zusammenfassung und Ausblick

pLOIN begrenzt die bereitgestellten Informationen und definiert, welche Informationen für einen spezifischen Anwendungsfall erforderlich sind. In der praktischen Anwendung sind Ansätze erforderlich, die eine Bereitstellung von Informationen ermöglichen, sodass diese bei Bedarf abgerufen werden können und die Entscheidungsfindung unterstützen. Mit Hilfe der heutigen technischen Möglichkeiten kann die Informationsbeschaffung und somit auch die Entscheidungsfindung vereinfacht werden. Jetzt müssen die entsprechenden Plattformen entwickelt werden und gleichzeitig die Arbeitsweisen sich so verändern, dass neuen Methoden ihr Potential voll entfalten können. Eine besondere Herausforderung wird dabei die Einbindung aller prozessbeteiligten Personen sein. Solange moderne Arbeitsweisen nicht angewendet werden und für die Projektdurchführung vorgeschrieben sind, wird es schwer, innovative Ideen in der Praxis umzusetzen.

Die Untersuchung dieser Vorgänge hat die Entwicklung eines firmeninternen pLOIN-Leitfadens angestoßen, auf den im Falle einer Anforderungen zurückgegriffen werden soll. So könnte für zukünftige interne Prozessbearbeitungen schneller geklärt werden, welche Informationen bei welcher Fragestellung zu erstellen bzw. zu überarbeiten sind. Es wird sich so auf die “wesentlichen” Informationen konzentriert werden, ohne zusätzliche Arbeiten auszuführen, die für den Entscheidungsprozess nicht von Relevanz sind. Ein unzureichendes Informations- und Datenmanagement kann zu Informationsverlusten führen, die den Lösungsfindungsprozess verzögern. Eine zentrale Ursache liegt darin, dass entscheidungsrelevante Informationen häufig nicht gezielt auffindbar sind. Dadurch besteht das Risiko, dass wesentliche Daten verloren gehen oder nicht rechtzeitig in den Entscheidungsprozess einfließen. Zwar werden die Informationen in einer Projekt-Cloud gespeichert, jedoch fehlt oft eine direkte Verknüpfung zu Begründungen oder Entscheidungsgrundlagen. Eine technische Lösung wird aktuell erarbeitet.

Literatur

[1] H.-R. Pfister, H. Jungermann und K. Fischer, *Die Psychologie der Entscheidung: Eine Einführung*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2017. DOI: 10.1007/978-3-662-53038-2

[2] A. Borrmann, M. König, C. Koch und J. Beetz, Hrsg., *Building Information Modeling: technologische Grundlagen und industrielle Praxis*, 2., aktualisierte Auflage, VDI-Buch, Wiesbaden: Springer Vieweg, 2021, 1 S. DOI: 10.1007/978-3-658-33361-4

[3] Karlheinz Pfarr, *Handbuch der kostenbewußten Bauplanung*. Deutscher Consulting Verlag Wuppertal, 1976, 320 S.

[4] G. Ropohl, *Eine Systemtheorie der Technik: zur Grundlegung der allgemeinen Technologie*. München Wien: Carl Hanser Verlag, 1979, 336 S.

[5] B. Kochendörfer, J. H. Liebchen und M. G. Viering, *Bau-Projekt-Management* (Leitfaden des Baubetriebs und der Bauwirtschaft). Wiesbaden: Springer Fachmedien Wiesbaden, 2018. DOI: 10.1007/978-3-8348-2245-1

[6] BIM Ruhr, Hrsg., *Digitalisierung der Baubranche: Stimmungsbild zu Building Information Modeling in kleinen und mittelständischen Betrieben*, 2021. besucht am 11. Juni 2025. Adresse: https://www.bim-ruhr.net/fileadmin/digitalisierung_in_der_baubranche_bim_ruhr.pdf

[7] *VDI/DIN-EE 2552 Blatt 12.1 Building Information Modeling - Struktur zu Beschreibung von BIM-Anwendungsfällen*, 2022. Adresse: www.vdi.de/2552

[8] DIN, *DIN EN ISO 7817-1 Bauwerksinformationsmodellierung Informationsbedarfstiefe Teil 1: Konzepte und Grundsätze (ISO 7817-1:2024); Deutsche Fassung EN ISO 7817-1:2024*, 2024.

[9] N. Mohan, G. Wolf, M. Wogan, J. Beilfuß und R. Groß, »Auf dem Weg zu einem ganzheitlichen Lebenszyklusmanagement von Bauwerken für nachhaltiges Bauen: Vorschlag einer Methodik zur digitalen Transformation durch datenbankbasierte Informationsmodelle«, in Ruhr-Universität Bochum, 2023. DOI: 10.13154/294-10090

[10] G. Wolf, M. Wogan, N. Mohan, N. Nejat, R. Gross und K. Menzel, »Advancing AEC project management: A model-based and data-driven approach for sustainable practices«, **presentedat**2024 European Conference on Computing in Construction, 2024. DOI: 10.35490/EC3.2024.198 Adresse: https://ec-3.org/publications/conference/paper/?id=EC32024_198

[11] M. Wogan, N. Mohan, G. Wolf, N. Nejat, K. Menzel und R. Gross, »Synergising lifecycle project management for sustainability: Towards a streamlined approach through different project phases«, **presentedat**The 9th International Conference on Civil, Structural and Transportation Engineering, 2024. DOI: 10.11159/iccste24.157 Adresse: https://avestia.com/ICCSTE2024_Proceedings/files/paper/ICCSTE_157.pdf

[12] M. Wogan, N. Mohan, K. Menzel, L. Trost und R. Gross, »Strategic advancements in AEC project management: Leveraging LOIN-based information models for total life cycle«, 2024.

[13] DIN, *DIN EN ISO 19650-1 Organisation und Digitalisierung von Informationen zu Bauwerken und Ingenieurleistungen, einschließlich Bauwerksinformationsmodellierung (BIM) Informationsmanagement mit BIM Teil 1: Begriffe und Grundsätze (ISO 19650-1:2018); Deutsche Fassung EN ISO 19650-1:2018*, Aug. 2019.

[14] T. Bachor, E. Straßenmeyer und F. Stevens. »Informationsanforderungen gemäß DIN EN ISO 19650«, besucht am 26. Feb. 2025. Adresse: https://www.build-ing.de/fachartikel/detail/informationsanforderungen-gemaess-dineniso19650/

[15] BIM4INFRA2020. »Steckbriefe der wichtigsten BIM-Anwendungsfälle«, besucht am 10. März 2025. Adresse: https://www.bimdeutschland.de/fileadmin/media/Downloads/Download-Liste/BIM4INFRA/3_6_BIM4INFRA2020_AP4_Teil6.pdf

[16] C. C. Eichler, C. Schranz, T. Krischmann, H. Urban, M. Hopferwieser und S. Fischer, *BIMcert Handbuch: Grundlagenwissen openBIM*, unter Mitarb. von TU Wien. Mironde-Verlag, 2024. DOI: 10.34726/5384 besucht am 21. Nov. 2024. Adresse: https://repositum.tuwien.at/handle/20.500.12708/192612

[17] H. Hamann, F. Petre und M. Wogan, »Berücksichtigung von Nutzeranforderungen bei der Planung von ÖPNV-Trassen: Eine Methode zur objektiven Bewertung und Empfehlungen im Kontext von Haltestellenstandorten«, unter Mitarb. von TUHH Universitätsbibliothek, 2024. DOI: 10.15480/882.13495

Building Knowledge Representation for Computer Scientists: Insights from Teaching

Diellza Elshani[1], Daniel Hernandez[2], Steffen Staab[2,3] and Thomas Wortmann[1]
[1]Institute for Computational Design and Construction, Department for Computing in Architecture (ICD/CA), University of Stuttgart, Germany
[2]Department of Analytic Computing, Institute for Artificial Intelligence, University of Stuttgart, Germany
[3]Electronics and Computer Science, University of Southampton, United Kingdom
E-mail(s): diellza.elshani@icd.uni-stuttgart.de, daniel.hernandez@ki.uni-stuttgart.de, steffen.staab@ipvs.uni-stuttgart.de, thomas.wortmann@icd.uni-stuttgart.de

Abstract: The Semantic Digital Twins Lab, held in winter 2023/24 at the University of Stuttgart, introduced computer science students to knowledge representation in the building industry. This paper highlights key insights and challenges, particularly in querying semantic and geometric data in building graphs. Students struggled with complex data structures and global coordinate computation—often beyond SPARQL's capabilities—but gained valuable experience and problem-solving skills. However, despite these challenges, the lab experience offered valuable learning opportunities and fostered innovative problem-solving skills among the students.

Keywords: Teaching, architectural modeling, geometric information, knowledge representation, semantic web technologies, IFC, logical inferences

DOI: 10.18154/RWTH-CONV-254895. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

The design, construction, and operation of a building compromise multiple disciplines and processes which must be performed simultaneously during the lifecycle of a building [1]. Giving digital objects semantics allows integrating knowledge across different disciplines, and ensuring a correct alignment. A comprehensive review of the application of semantic web technologies in BIM highlights both the potential and current limitations of these approaches in AEC contexts [2]. Recent work has demonstrated how semantic technologies can support co-design workflows and enable cross-domain querying and reasoning in AEC settings [3], reinforcing the educational value of introducing these concepts through practical lab environments. A research project of the Cluster of Excellence Integrative Computational Design and Construction for Architecture (IntCDC) at the University of Stuttgart, tackles these challenges. This research project is a joint effort between two institutes, the Department of Analytic Computing, Institute for Artificial Intelligence (KI/AC) and the Institute for Computational Design and Construction, Department for Computing in Architecture (ICD/CA). KI/AC teaches a course on Knowledge Graphs, which presents the fundamentals of the Semantic Web. Although this course

combines lectures and exercises, it does not show their application in a particular use case or industry. At the same time, ICD/CA is involved in teaching the Building Information Modelling (BIM) seminar, which teaches several tools and practices of the building industry for creating, collaborating and exchanging information, but does not cover semantic web technologies in-depth. We thus desinged a course where students practice Semantic Web technologies in the building industry. This new course was conducted during the winter semester of 2023/24. The course included lectures and tasks covering semantic web technologies [4] and knowledge representation pertinent to the building sector [5], [6], [7]. Ten students from diverse Master's programs—including Computer Science, Infotech, Informatics, and Autonomous Systems—took part in the course. We utilized the open, international standard (ISO 16739-1:2018) data schema that supports the exchange of digital data among heterogeneous applications, the Industry Foundation Classes (IFC) [8]. Students used several IFC files following the IFC data schema, including the duplex house (including several disciplinary representations) and structural models in an IFC format from the Semantic Enrichment challenge of the LDAC2022 Challenge[1] [9]. In the world of semantic web technologies, students mainly used the Resource Description Framework (RDF), the Web Ontology Language (OWL) [10], and the query language SPARQL. They worked in groups, they undertook tasks ranging from converting IFC files to RDF (LBD ontologies, ifcOWL [11], and developing their own parser), and visualizing 3D models and RDF graphs to integrating features into a web application with data loading and querying capabilities, leveraging GraphDB[2]. Weekly sessions focused on group work, progress updates, and applying newly learned skills. The course concluded with tasks involving logical inference from semantic and geometric data, such as detecting pipe intersections and identifying structural supports. This paper examines the use of semantic web technologies in architectural modeling education, focusing on BIM data integration, knowledge graphs, and linking data from different disciplines.

2 Lab Structure

The Semantic Digital Twins Lab was organized into weekly tasks and a final comprehensive task involving semantic and geometric querying on graph-based building models. The lab aimed to introduce students to semantic web technologies and knowledge representation in the building industry. Learning objectives included mastering RDF conversion, querying geometric data, applying ontologies, and solving problems related to building information. Through hands-on and collaborative exercises, students gained experience in representing and managing architectural data using semantic web principles, with a focus on real-world applicability. Each week, students worked on tasks aligned with specific Intended Learning Outcomes (ILOs), covering skills such as defining software architecture, visualizing graphs, converting IFC files to RDF, using ontologies like ifcOWL and LBD, integrating systems, and building web interfaces. Students used different software, including the Java implementation of the Ruby language (JRuby), the Ruby RDF library, the Python RDF library RDFlib, the Apache Jena library for RDF file parsing, the GraphDB database for loading and querying RDF data with SPARQL, the Java tool for IFC to RDF conversion, the IfcOpenShell tool to facilitate IFC file reading and extracting relevant information, the D3 JavaScript library to visualize RDF graphs, and

[1] https://github.com/SSoLDAC-2022/Challenge-4
[2] GraphDB is a graph database and knowledge discovery tool compliant with RDF and SPARQL

Protégé for ontology exploration. Table 1 outlines these ILOs and their corresponding tasks. The final task (ILO#14) required each group to tackle a complex problem in building industry knowledge representation. This involved integrating data across disciplines, exploring building elements, interpreting geometry, and understanding implicit vs. explicit information in models. Students deepened their proficiency in SPARQL, RDF, ontologies, and geometric representation. The following subsections present the four core tasks tackled by student groups, highlighting both outcomes and challenges.

Table 1: Intended Learning Outcomes (ILOs) of the Lab and Corresponding Tasks

ILO	Definition	Implementation Task
ILO #1	Develop the ability to define and conceptualize software architecture within the context of building industry knowledge representation.	Exploratory task for each group: define a software architecture.
ILO #2	Acquire the skills to integrate software components and systems within the context of building industry knowledge representation.	Software integration.
ILO #3	Improve graph understanding through visualization and interpreting building model representations.	Visualize graph.
ILO #4	Gain proficiency in converting IFC files to RDF format.	Convert IFC files to RDF graphs that do not follow a certain ontology.
ILO #5	Develop the ability to create a web interface for uploading IFC files, a crucial aspect of data management in the building industry.	Create a webform where users can upload IFC files.
ILO #6	Understand building models through geometry, thereby acquiring the skills necessary to visualize and interpret simple geometry on web platforms.	Visualize IFC geometry on the web (example geometry: Box).
ILO #7	Gain proficiency in parsing IFC files into existing ontologies, such as the Linked Building Data (LBD) ontologies.	Convert IFC files into LBD ontologies.
ILO #8	Develop the capability to store IFC and RDF files in a Graph Database, essential for data management and querying in the building industry.	Pushing the IFC and RDF files to Graph DB.
ILO #9	Understand building models through geometry, thereby acquiring the skills necessary to visualize and interpret IFC geometry on web platforms.	Visualize IFC file.
ILO #10	Gain proficiency in parsing IFC files into existing ontologies, such as ifcOWL formats.	Convert IFC files into ifcOWL graphs.
ILO #11	Develop the ability to define axioms within the context of building industry knowledge representation.	Defining Axioms, Exploratory task.
ILO #12	Acquire skills in inferring axioms, valuable for logical inferences in the building industry.	Inferring Axioms.
ILO #13	Gain proficiency in inferring facts and validating relationships using ontology editor software.	Infer facts using Protégé.
ILO #14	Learn data integration with semantic web technologies, by combining two disciplinary models for the same building.	Adding supports relation between architectural elements; adding connectsTo relation between transit spaces; linking pipes with architectural elements; linking electrical appliances with spaces.

2.1 Task 1: Adding “Supports” Relation between Architectural Elements

Task description: This task focused on inferring structural support relationships between architectural elements in the building model. Since these relations were not explicitly included in the IFC-derived ifcOWL data, students had to identify which elements likely support others—such as columns supporting beams—and represent this connection using a new predicate, "supports". As the information was missing from the model, students were required to make assumptions and define heuristics to infer the relation. After adding this relation to the mode, the students must answer the following questions: (1) Which elements are not supported? (2) Which elements are supported by a specific element? (2) Which classes include elements that support other elements? (3) Which classes include elements that do not support other elements? **Students' Solution**: The students assumed that the “support” relation can be inferred from the position of the building elements. If a building element is directly above another building element, then the students assumed that the element below is supporting

the element above. To simplify the operation with the irregular geometries of objects, the students assumed that this relationship can be inferred from the bounding boxes of elements. To reduce the search space, the students assumed that only certain types of building elements are designed to support other elements, the elements that have the attribute load bearing with the value true. They calculate the bounding box by finding the maximum and minimum values for the element coordinates regarding a global reference system. Hence, they translated these coordinates to a global coordinate system before computing the bounding boxes. They did this recursively for elements whose position is relative to a position of another element. Finally, they added the relation defined according to these assumptions, and they formulated the requested queries using SPARQL.

Evaluation of the learning objectives: Although the students completed the task, the goal of practicing with the fundamental concepts of ontologies was not met. In particular, they do not use the notion of inheritance of classes. Instead of finding arbitrary elements which support other elements, they find elements of particular classes (e.g., Column). Instead of exploring the ontology at a high-level, they explore patterns in the relationships between individuals of lower-level classes. Besides the planned objectives, they learned about the representation of coordinated systems.

2.2 Task 2: Adding "ConnectsTo" Relation between Spaces

Task Description: This task involved inferring spatial connectivity between rooms in the building model. While the IFC-derived data defined enclosed spaces, it did not explicitly state whether movement was possible between them. Students were asked to add this missing information by introducing a symmetric predicate, "connectsTo", between spaces that are directly connected. This connection could be inferred from the spatial geometry—either through openings like doors or direct adjacency without intervening elements. After adding this relation, students had to test if the relation was successfully incorporated by querying the SPARQL endpoint with the following queries: (1) Do we need to cross space A when transiting from space B to space C?, (2) Which spaces can be reached from space A without passing through another space?. **Students' Solution**: The students' approach was to utilize IFC opening elements and identify spaces related to the same ifcOpeningElement predicate. Figure 1 shows the steps the group took to tackle the task.

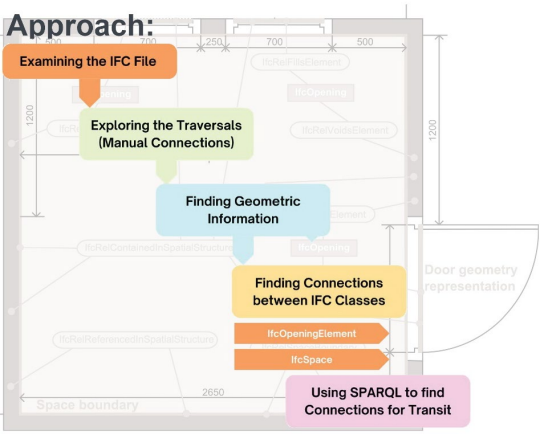


Figure 1: Approach of students in adding "ConnectsTo" Relation between Spaces.

Evaluation of Learning Objectives: Students completed the task by adding the "connectsTo" relation between some spaces. However, they did not succeed in correctly adding this relation, because they added the relation between spaces that are connected by an ifcOpeningElement. An ifcOpeningElement, in essence, can also be a non-transit element, such as a window. Nevertheless, they gained insights into spatial relationships and navigation pathways by analyzing the hierarchy of the ontology classes, as well as the physical representation of the building.

2.3 Task 3: Linking Pipes with Architectural Elements

Task Description: This task required students to integrate separate disciplinary models by inferring spatial relationships between pipes and architectural elements. Specifically, they had to identify where pipes intersect with building components, such as walls, and make this relation explicit using the predicate crosses. The intersection information had to be derived from the geometric data provided in the mechanical, plumbing, and architectural models. Finally, students were asked to test the success of incorporating this relation by answering the following queries: (1) Does any pipe cross a load-bearing wall? (2) Which walls are crossed by pipes? **Students' Solution**: Despite encountering two primary challenges—understanding the class hierarchy of IFC files and extracting global coordinates of objects—the student group successfully addressed the task. They manually explored the graph and utilized GraphDB to deepen their understanding of the data structures involved. Employing IfcOpenShell facilitated the retrieval of relevant information from the IFC, allowing them to insert it into the graph using SPARQL INSERT DATA queries. Subsequently, they crafted queries to retrieve coordinates of walls and pipes, followed by calculations of bounding boxes and intersections using SPARQL queries (see Algorithm 1, Algorithm 2, and Algorithm 3).

Evaluation of Learning Objectives: While successfully linking pipes with architectural elements, students engaged in manual exploration and utilized tools like GraphDB and IfcOpenShell to overcome challenges. This hands-on approach fostered a deeper understanding of data structures and practical application of SPARQL queries.

SPARQL Algorithm: Coordinates of Wall Geometry from IFC File

Algorithm 1: Query to retrieve the coordinates of walls (e.g., from an IFC architecture file).

```
PREFIX ifc:     <http://standards.buildingsmart.org/IFC/DEV/IFC2x3/TC1/OWL#>
PREFIX wall:    <http://linkedbuildingdata.net/ifc/resources20240127_210211/IfcWallStandardCase_>
PREFIX list:    <https://w3id.org/list#>
PREFIX express:<https://w3id.org/express#>
PREFIX rdf:     <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

SELECT DISTINCT ?wall ?Polyline ?CartesianPointList ?Point ?Point2 ?Point3 ?Point4 ?Point5
FROM <https://ex.org/duplex_pieter>
WHERE {
  ?wall a ifc:IfcWallStandardCase;  ifc:representation_IfcProduct ?ProductShape .
  ?ProductShape ifc:representations_IfcProductRepresentation ?RepresentationList .
  ?RepresentationList list:hasContents ?ShapeRepresentation .
  ?ShapeRepresentation ifc:items_IfcRepresentation ?Polyline .
  ?Polyline ifc:points_IfcPolyline ?CartesianPointList .

  ?CartesianPointList list:hasContents ?Point ;
                      list:hasNext/list:hasContents ?Point2 ;
                      list:hasNext/list:hasNext/list:hasContents ?Point3 ;
                      list:hasNext/list:hasNext/list:hasNext/list:hasContents ?Point4 ;
                      list:hasNext/list:hasNext/list:hasNext/list:hasNext ?Point5 .
}
```

SPARQL Algorithm: Calculating Bounding Box for Pipes

Algorithm 2: Calculating pipe bounding box based on position, radius, and depth.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX ifc: <http://www.buildingsmart-tech.org/ifcOWL/IFC4#>
PREFIX ex:  <http://example.org/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

SELECT ?wall
WHERE {
  GRAPH <https://ex.org/new_pipe_data> {
    ?pipe a ex:IfcFlowSegment;  ex:x ?pipeX; ex:y ?pipeY; ex:z ?pipeZ;  ex:radius ?radius; ex:depth ?pipeDepth .
    BIND (xsd:double(?pipeX) - xsd:double(?radius)       AS ?pipeMinX)
    BIND (xsd:double(?pipeY) - xsd:double(?radius)       AS ?pipeMinY)
    BIND (xsd:double(?pipeZ)                             AS ?pipeMinZ)
    BIND (xsd:double(?pipeX) + xsd:double(?radius)       AS ?pipeMaxX)
    BIND (xsd:double(?pipeY) + xsd:double(?radius)       AS ?pipeMaxY)
    BIND (xsd:double(?pipeZ) + xsd:double(?pipeDepth)    AS ?pipeMaxZ)
    FILTER (
      (?wallMaxX >= ?pipeMinX && ?wallMinX <= ?pipeMaxX) ||
      (?wallMaxY >= ?pipeMinY && ?wallMinY <= ?pipeMaxY) ||
      (?wallMaxZ >= ?pipeMinZ && ?wallMinZ <= ?pipeMaxZ)
    )
  }
}
```

SPARQL Algorithm: Inserting Intersection Data and Querying Crossings

Algorithm 3: Inserting new intersection triples and querying out :crossesWith relations.

```
PREFIX ifc:  <http://standards.buildingsmart.org/IFC/DEV/IFC2X3/TC1/OWL#>
PREFIX rdf:  <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX inst: <http://linkedbuildingdata.net/ifc/resources20231206_102530/>

INSERT {
  GRAPH <https://ex.org/intersections> {
    inst:Pipe_54976 ifc:crossesWith inst:Wall_4131 ;
                    ifc:crossesWith inst:Wall_4287 ;
                    ifc:crossesWith inst:Wall_4399 ;
                    ifc:crossesWith inst:Wall_4465 ;
                    ifc:crossesWith inst:Wall_5448 .
  }
}
SELECT ?wall
WHERE {
  ?pipe rdf:type ifc:IfcFlowSegment;  ifc:crossesWith ?wall.
  ?wall rdf:type ifc:IFCWallStandardCase .
}
```

2.4 Task 4: Linking Electrical Appliances with Spaces

Task Description: This task involved integrating electrical and architectural models by inferring the spatial location of electrical appliances within the building. Students were asked to establish explicit connections between appliances and the spaces they occupy using the predicate locatedIn. Since this spatial relation was not directly available in the IFC data, it had to be inferred from the geometric information. To validate the integration, students queried the followings: (1) Which room has the most electrical appliances?, (2) Which two appliances are very close to each other? (so that we can use the same circuit). **Students' Solution**: Students initially visualized the IFC files to comprehend their structure. They attempted to establish a linkage between the files but encountered challenges in aligning spaces. Additionally, they implemented functionality to find the closest appliances, but faced difficulties with SPARQL while integrating the electrical file with the architectural file due to disparities in space identification and the unavailability of global coordinates for the architectural file. These obstacles hindered the integration process.

Evaluation of Learning Objectives: Despite completing the task for some elements, the students faced challenges in determining coordinates for all rooms. Specifically, they struggled with computing the global coordinates of geometric representations due to the intricate nature of local placements. However, through these challenges, they gained a deeper understanding of building models. They still managed to implemented the process of integrating two IFC files and delved into understanding the coordinate system, both relative and global.

3 Conclusion

The Semantic Digital Twins Lab introduced students to knowledge representation challenges within the building industry through practical, interdisciplinary tasks. The course offered hands-on experience with structured data modeling, querying, and inference in the context of digital building representations. The most significant challenges arose from understanding how building models are encoded using ifcOWL and from extracting topological and spatial relationships between building components. These geometries are often represented in local coordinate systems, requiring recursive transformations to derive global positions. Students found this process particularly demanding, as SPARQL offers only limited recursion capabilities and lacks support for non-commutative transformations such as rotations. We observed that these technical hurdles risk overshadowing the broader learning objectives of semantic modeling. To address this, future iterations of the lab will place greater emphasis on modeling fundamentals from the outset. Specifically, we plan to (i) enrich the current ifcOWL datasets with concepts from the BOT ontology to improve semantic clarity and (ii) introduce geometric encoding and coordinate transformation principles earlier in the course. The course was attended primarily by students from computer science and related disciplines. While they demonstrated strong proficiency in programming, RDF modeling, and database querying, difficulties with spatial reasoning and geometric interpretation highlighted the absence of architectural perspectives. In the future, we aim to include students from architectural and design backgrounds to better support interdisciplinary collaboration and domain-relevant problem-solving. Students responded positively to the structure of the lab. Working in groups fostered collaboration and knowledge exchange, while the interactive format—with regular feedback and integration between lectures and hands-on tasks—was well received. The diversity of tutor backgrounds, spanning computer science, AI, and architecture, further enriched the learning experience and was noted as beneficial in student feedback. These findings support our conviction that an interactive, interdisciplinary setup is essential for effectively teaching the application of semantic web technologies in the AEC domain. Overall, the lab demonstrated both the potential and the current limitations of integrating semantic web technologies in AEC education, offering valuable lessons for curriculum development at the intersection of BIM and knowledge representation.

4 Acknowledgements

Supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2120/1 – 390831618 and COFFEE - STA 572_15-2. The authors thank all course participants: Ankita Saha, Annie Bhalla, Chunzhi Guo, Haomin Shi, Long Bai, Mayank Jha, Robin Schalkowsky, Shivani Soni, Varun Bhat, and Yao Chu.

References

[1] D. Elshani, T. Wortmann, and S. Staab, “Towards Better Co-Design with Disciplinary Ontologies: Review and Evaluation of Data Interoperability in the AEC Industry”, in *Proceedings of the 10th Linked Data in Architecture and Construction Workshop (LDAC 2022)*, vol. Vol-3213, Hersonissos, Greece: CEUR Workshop Proceedings, 2022, pp. 43–52.

[2] X. Zhang, J. Beetz, and M. Weise, “A Review of the Application of Semantic Web Technologies in Building Information Modeling”, *Archives of Computational Methods in Engineering*, vol. 29, pp. 1991–2022, 2022. DOI: 10.1007/s11831-021-09595-6

[3] D. Elshani, A. Lombardi, D. Hernandez, S. Staab, A. Fisher, and T. Wortmann, “AEC Co-Design Workflow for Cross-Domain Querying and Reasoning Using Semantic Web Technologies”, *Automation in Construction*, vol. 176, p. 106 226, 2025. DOI: 10.1016/j.autcon.2025.106226

[4] A. Hogan et al., *Knowledge Graphs*. New York, NY, USA: Association for Computing Machinery, Jul. 2021, vol. Vol-54. DOI: 10.1145/3447772

[5] D. Elshani et al., “Building Information Validation and Reasoning Using Semantic Web Technologies”, en, in *Computer-Aided Architectural Design. INTERCONNECTIONS: Co-Computing Beyond Boundaries*, vol. 1819, Cham: Springer Nature Switzerland, 2023, pp. 470–484. DOI: 10.1007/978-3-031-37189-9_31

[6] P. Pauwels and K. McGlinn, editors, *Buildings and Semantics: Data Models and Web Technologies for the Built Environment*, First edition. Leiden, The Netherlands ; Boca Raton, FL: CRC Press/Balkema, 2023. DOI: 10.1201/9781003204381

[7] D. Elshani, A. Lombardi, A. Fisher, D. Hernandez, S. Staab, and T. Wortmann, “Knowledge Graphs for Multidisciplinary Co-Design: Introducing RDF to BHoM”, in *Proceedings of the 10th Linked Data in Architecture and Construction Workshop (LDAC 2022)*, vol. Vol-3213, Hersonissos, Greece: CEUR Workshop Proceedings, 2022, pp. 32–42. [Online]. Available: http://ceur-ws.org/Vol-3213/

[8] buildingSMART International, *Industry Foundation Classes (IFC)*, buildingSMART International, 2018. Accessed: Oct. 10, 2024. [Online]. Available: https://www.buildingsmart.org/standards/bsi-standards/industry-foundation-classes/

[9] T. Bloch and R. Sacks, “Clustering Information Types for Semantic Enrichment of Building Information Models to Support Automated Code Compliance Checking”, *Journal of Computing in Civil Engineering*, vol. 34, no. 6, p. 04 020 040, 2020. DOI: 10.1061/(ASCE)CP.1943-5487.0000922

[10] P. Hitzler, M. Krötzsch, B. Parsia, P. F. Patel-Schneider, and S. Rudolph, *OWL 2 Web Ontology Language Primer (Second Edition). W3C Recommendation*, https://www.w3.org/TR/owl2-primer/.

[11] P. Pauwels and W. Terkaj, “EXPRESS to OWL for Construction Industry: Towards a Recommendable and Usable ifcOWL Ontology”, *Automation in Construction*, vol. 63, pp. 100–133, 2016. DOI: https://doi.org/10.1016/j.autcon.2015.12.003

Integrated Real-Time Monitoring of Sewer Systems via SPC and Graph-Based Validation

Jun Li[1] and Finia Buitkamp[2]
[1]Geodetic Institute and Chair of Computing in Civil Engineering & GIS (gia), Aachen, Germany
[2]Wasser-Abwasser Systemtechnik GmbH (W.A.S.)
E-mail(s): jun.li@gia.rwth-aachen.de, buitkamp@was-bs.de

Abstract: Urban sewer systems face growing challenges due to aging infrastructure, climate-induced stress, and sensor unreliability. This paper presents a hybrid framework that combines Statistical Process Control (SPC) with graph-based spatial validation to monitor sewer networks in real time. Temporal anomalies are identified through SPC techniques - including Western Electric rules, CUSUM, EWMA - applied to streaming sensor data. To enhance detection performance under varying hydrological conditions, the framework incorporates weather-based classification prior to SPC computation. Simultaneously, flow consistency across the network is validated using a directed graph model, detecting structural inconsistencies and sensor faults. Experimental results using real-world data demonstrate that SPC methods capture local anomalies effectively, while spatial validation uncovers topological violations that temporal methods may miss. The combination of these methods enables robust, real-time detection of flow disturbances and supports infrastructure monitoring in smart sewer systems.

Keywords: real-time monitoring, sewer networks, anomaly detection, SPC, spatial validation

DOI: 10.18154/RWTH-CONV-254896. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Urban drainage systems are increasingly stressed by aging infrastructure and climate-induced extremes. In Germany, annual precipitation has risen by approximately 7% since 1881, and the frequency of localized heavy rainfall events has significantly increased [1], [2]. Combined sewer systems (CSS), still prevalent in many cities, are often overwhelmed by stormwater during intense rainfall, resulting in untreated combined sewer overflows (CSOs) into surface waters [3].

Beyond hydraulic capacity issues, modern urban sewer systems rely on a dense sensor network to monitor flow rates, water levels, and other hydrodynamic parameters. However, these sensors are prone to noise, drift, and measurement bias [4], [5]. The resulting data quality challenges complicate early-warning detection tasks. Traditional anomaly detection in such systems still involves labor-intensive visual inspections or heuristic thresholds, limiting the scalability of real-time monitoring [6], [7].

To address these challenges, data-driven methods such as Statistical Process Control (SPC) and graph-based validation frameworks are gaining attention for their ability to support automated, scalable, and interpretable anomaly detection [5], [8].

The research project "Kanalanalyse", a collaboration between xxx and xxx, addresses this gap by integrating real-time sensor data with historical and simulated records. This paper presents two key modules of the system:

(1) A temporal anomaly detection component using Statistical Process Control (SPC) methods, structured in two phases - baseline modeling and real-time monitoring - applied to sewer flow and level data;

(2) A spatial consistency check based on directed graph topology, which validates flow monotonicity to detect implausible measurements and sensor faults.

Recent studies emphasize the potential of data-driven methods to enhance operational awareness in sewer networks, particularly under dynamic weather conditions [5]. To support this, rainfall events are further characterized based on their temporal structure and intensity, enabling the differentiation between convective and stratiform patterns. These contextual insights guide the interpretation of flow anomalies, allowing the system to distinguish between hydraulic responses to actual events and deviations caused by sensor faults.

Together, these methods enable a scalable, real-time monitoring framework that supports early warning and improves data quality across urban drainage networks.

2 Spatial Validation of Flow Consistency

Urban sewer systems can be broadly categorized into separate and combined systems. In combined systems, both sanitary sewage and stormwater runoff are conveyed in the same pipes, leading to potential overflows during heavy rain events.

Validating flow consistency requires understanding flow direction and upstream-downstream relations. We model the sewer network as a directed graph and check spatial consistency via flow monotonicity constraints.

2.1 Graph-Based Modeling and Flow Monotonicity Validation

Sewer networks with junctions, pipes, and manholes map naturally to directed graphs $G = (V, E)$, where V are junctions or sensor locations and E are conduits aligned with flow direction. This structure supports efficient modeling of hydraulic dependencies and topology-aware analysis.

Directed graphs have been widely used in sewer system design and monitoring. For optimization, they support layout simplification and hydraulic simulations [9], while in monitoring applications, they capture flow dependencies for spatio-temporal analysis [4], [10].

Building on this structure, we introduce a physically motivated validation method based on flow monotonicity: under normal conditions, the cumulative flow from upstream nodes should not exceed that at the downstream node, aside from minor deviations due to measurement noise or unmodeled inflows. Formally, for a downstream node v and its set of upstream neighbors $\mathcal{U}(v)$, we expect:

$$\sum_{u \in \mathcal{U}(v)} Q_u(t) \leq Q_v(t) + \epsilon, \tag{1}$$

where $Q_u(t)$ and $Q_v(t)$ are flow measurements and ϵ accounts for sensor uncertainty and spatial delays. Systematic violations of this inequality may indicate sensor faults, data desynchronization, undocumented inflows, or structural inconsistencies. We apply this constraint across all connected node pairs to validate the spatial coherence of sensor data in real-time.

2.2 Algorithm: Upstream–Downstream Consistency Check

Based on the monotonicity constraint introduced in Section 2, we implement a spatial validation procedure that systematically checks for flow inconsistencies across upstream-downstream sensor pairs in the sewer network.

Let $G = (V, E)$ be the directed graph representing the sewer network, where each node $v \in V$ is associated with a time series $Q_v(t)$ denoting the measured flow at time t. For each node v, let $\mathcal{U}(v)$ denote the set of its direct upstream neighbors (i.e., $(u, v) \in E$). For each time step t, the algorithm performs the following:

1. For each node $v \in V$ with incoming edges:
   (a) Retrieve $Q_v(t)$ and $\{Q_u(t) \mid u \in \mathcal{U}(v)\}$.
   (b) Compute the difference $\Delta_v(t) = Q_v(t) - \sum_{u \in \mathcal{U}(v)} Q_u(t)$.
   (c) If $\Delta_v(t) < -\epsilon$, flag this as a monotonicity violation.
2. Store all violations $(t, u, v, \Delta_v(t))$ for reporting or visualization.

The tolerance term ϵ is set as a fixed violation rate of 10% of the downstream flow at time t, i.e., $\epsilon = 0.1 \cdot Q_v(t)$. This accounts for measurement uncertainty, short-term flow fluctuations, and minor unrecorded inflows.

To facilitate practical application, the consistency check algorithm has been integrated into a stream-processing pipeline and deployed as part of a real-time monitoring system developed in the research project "Kanalanalyse". Detected violations are streamed to a web-based dashboard, enabling operators to visually inspect the spatial distribution and temporal evolution of flow anomalies in the sewer network.

3 Statistical Process Control for Sewer Flow Monitoring

Sewer systems are increasingly equipped with flow and level sensors that enable continuous monitoring. However, interpreting this time series data remains challenging due to its susceptibility to seasonal variations, rainfall events, sensor drift, and infrastructure anomalies. Statistical Process Control (SPC) offers a framework originally developed for industrial quality control that has proven effective in environmental and infrastructure monitoring [8], [11].

SPC distinguishes between normal process variability and deviations that may indicate underlying issues. In the context of wastewater networks, SPC has been applied for detecting anomalies such as blockages, overflows, and sensor failures [6]. It has also been used for burst detection in

water distribution systems [11], flood risk index construction based on control limits [12], and overflow behavior analysis under different rainfall intensities [13]. Its application to sewer monitoring is motivated by its interpretability, computational efficiency, and suitability for real-time deployment.

3.1 SPC Methods

SPC methods rely on control charts to distinguish random variability from significant process deviations. Each chart consists of a center line representing the process average and upper/lower control limits that define acceptable fluctuation ranges. As illustrated in 1, When observations fall outside the control limits, they are flagged as potential anomalies. The example shows a Shewhart control chart with center line, upper control limit (UCL), and lower control limit (LCL), highlighting out-of-control points.

It is been proven that 6 sigma is a very effective framework for data quality control [14] This range captures approximately 99.73% of all data points under normal operation, thus providing a statistically justified threshold for anomaly detection [8].

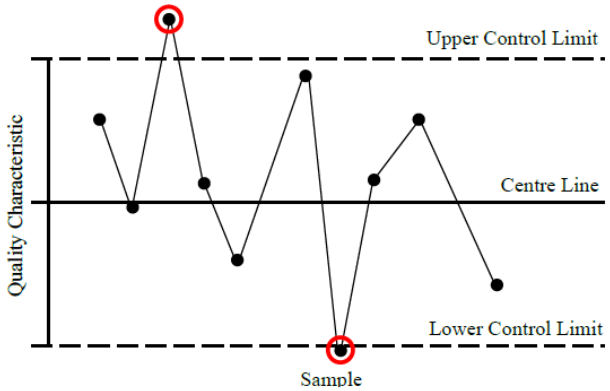


Figure 1: Example Shewhart Control Chart showing centre line, UCL, and LCL. Out-of-control points are highlighted.

Among SPC techniques, **the Western Electric Company (WEC) rules** represent a foundational approach. The first and most basic rule states that a process is considered out of control if any single data point falls outside the three-sigma range $[\mu - 3\sigma, \mu + 3\sigma]$. In this study, we adopt this rule by directly setting the control chart bounds to this interval. This provides a balance between sensitivity and false alarm rate in high-frequency environmental monitoring scenarios, and serves as a baseline against which more advanced methods (e.g., CUSUM and EWMA) are later compared.

EWMA (Exponentially Weighted Moving Average): This method assigns more weight to recent observations, making it effective for detecting gradual trends or small drifts in the data [8]. The control statistic is recursively updated as $z_t = \lambda x_t + (1 - \lambda) z_{t-1}$, where $\lambda \in (0, 1]$ is the smoothing factor. Control limits are calculated as $\mu \pm L \cdot \sigma \sqrt{\frac{\lambda}{2-\lambda} \cdot (1 - (1 - \lambda)^{2t})}$, based on the segment-specific mean μ and standard deviation σ.

CUSUM (Cumulative Sum Control Chart): A sequential analysis technique designed for identifying small, sustained shifts in the process mean [11]. It accumulates deviations from the target mean over time and flags anomalies when the cumulative statistic crosses a decision interval. In our

implementation, the CUSUM statistic is computed as the cumulative sum of deviations from the segment-specific mean μ.

3.2 SPC Phases and Implementation

Following standard SPC methodology [8], our framework comprises two phases:

Phase I — Baseline modeling: Historical data under stable conditions is used to compute process statistics - mean μ, standard deviation σ. Control limits (e.g., $\mu \pm 3\sigma$) are defined per method. Periods with anomalies or strong seasonality are excluded.

Phase II — Online monitoring: Incoming sensor data is checked against these limits. Violations - such as boundary crossings or sensitizing patterns (e.g., WEC rules) - trigger anomaly flags.

We implemented this two-phase approach in a real-time stream processing setup. Sensor data is transmitted using MQTT (Message Queuing Telemetry Transport), a lightweight publish-subscribe protocol optimized for reliable delivery in low-bandwidth environments.

The processing pipeline includes:

- **Preprocessing:** hourly resampling, gap filling, and outlier filtering.
- **Dynamic SPC:** incremental updates of CUSUM and EWMA per new data point.
- **Alerting:** a 10% violation threshold within a time window triggers alerts, filtering transient noise.

This architecture enables scalable integration of SPC into live sewer monitoring systems.

3.3 Experimental Setup

This section investigates whether SPC methods can reliably detect process anomalies in wastewater level data and how weather classification affects their interpretability. We evaluate three univariate SPC methods: WEC rules, EWMA, and CUSUM. For EWMA, we adopt Montgomery's recommended setting of a smoothing factor $\lambda = 0.2$ and control width $L = 3$, which balances sensitivity and false alarms [14]. [14] notes that $\lambda \in [0.05, 0.25]$ is commonly used, with $L = 3$ working well, especially for larger λ. For CUSUM, we use the standard decision interval $H = 5\sigma$, which reliably detects moderate persistent shifts with low false-alarm rates [14]. These parameter choices follow established practice and ensure fair comparability across methods.

Since rainfall strongly affects sewer behavior, especially in runoff-sensitive lines, we follow prior studies [5], [6], [13] in applying weather classification before SPC. Rainfall introduces transient variability that violates the stationarity assumptions of SPC. To address this, we divide the monitoring period into three meteorologically homogeneous segments based on a practical rainfall threshold of 0.3 mm per 15-minute interval: a dry weather period, a rain-affected period, and a post-event recovery phase.

To illustrate this need, we first applied SPC methods to the unsegmented series. When SPC methods were applied directly to the full time series, both WEC and EWMA exhibited strong responsiveness to rainfall dynamics (Figure 2 b)-d)). Each alarmed segment corresponded precisely to precipitation peaks (Figure 2 a)), suggesting that these methods are well-suited for early warning of potential overflow risks under storm conditions. In contrast, CUSUM produced ambiguous results. Since the full monitoring window spans dry, rainy, and post-event conditions, the use of a single global mean and

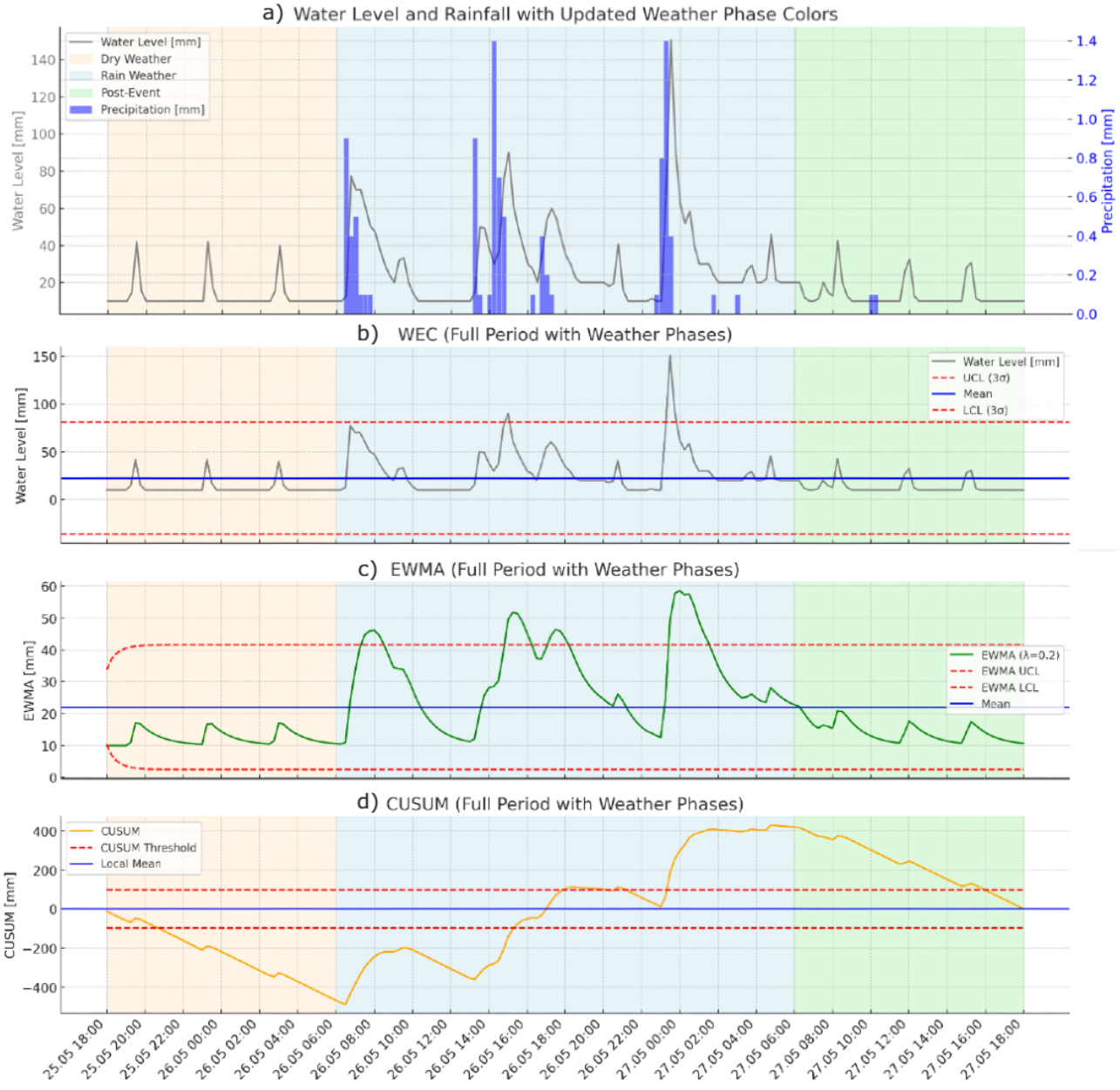


Figure 2: Rainfall with annotated weather phases and SPC output without weather classification.

standard deviation distorted its anomaly detection capability. CUSUM failed to distinguish between genuine system anomalies and statistical shifts caused by changing weather states. This highlights that, in order to accurately detect localized system anomalies, it is essential to first perform weather classification to eliminate weather-induced variability.

We then reapplied the SPC methods to the segmented data, distinguishing between dry weather and rain-affected periods to represent baseline and disturbed system states. The results confirmed statistical stability for all SPC methods, with no control limit violations detected during dry weather days. In contrast, significant deviations associated with precipitation events were identified in the rain affected segments. Notably, the CUSUM method, as illustrated in Figure 3, effectively addressed the issue of false cumulative signals caused by precipitation, which previously led to misinterpretation. It also delivered performance comparable to the other two methods. These findings demonstrate that applying weather classification prior to SPC significantly enhances interpretability and detection accuracy in real world sewer monitoring scenarios. Although our evaluation focuses on a single measurement site, the proposed stratified SPC approach is readily transferable to other locations within urban drainage networks where flow behavior is sensitive to weather conditions.

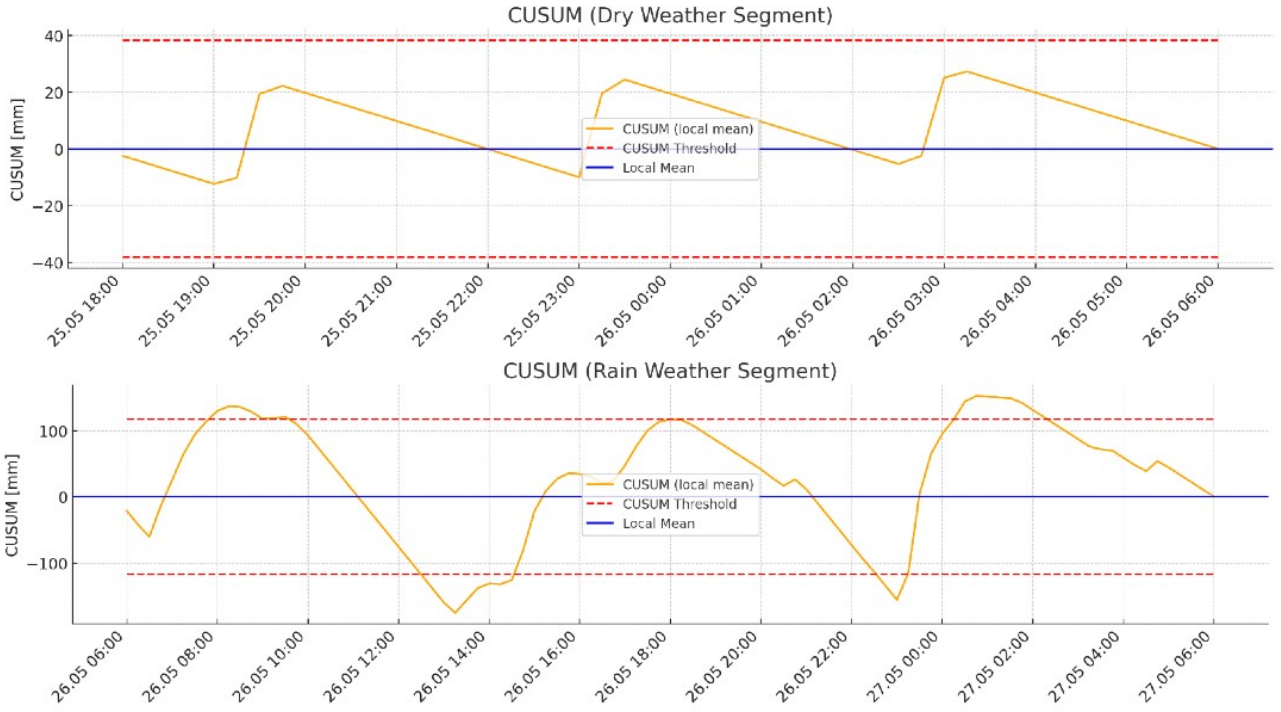


Figure 3: For the dry weather segment, the CUSUM method indicates statistical control; for the rain affected segment, it detects significant deviations associated with precipitation events.

4 Summary and Conclusions

We evaluated three SPC methods for temporal anomaly detection in sewer monitoring and found that applying them to unsegmented time series often produced misleading results, especially for CUSUM. In contrast, WEC and EWMA remained reliable under global conditions. When rainfall-based segmentation was introduced, all methods, including CUSUM, showed improved interpretability

and localization of anomalies. Complementing temporal analysis, we introduced a spatial validation approach based on flow-directed graph topology to detect violations of flow monotonicity. These violations can indicate backwater effects, undocumented inflows, or potential sensor inconsistencies - capturing anomalies that temporal methods alone might miss. Together, our findings highlight the necessity of weather classification and spatial validation as preprocessing steps for robust sewer monitoring. SPC should serve as a fast, interpretable pre-alarm system, while further diagnosis should involve physical modeling. Future work will explore adaptive time windows and integrate SPC with hydraulic simulations in a multi-stage monitoring architecture.

References

[1] IPCC, *Climate change 2022: Impacts, adaptation and vulnerability. working group ii contribution to the sixth assessment report*, Online, 2022.

[2] Umweltbundesamt, *Klimawandel und urbane sturzfluten: Risiken für städte*, Online, 2022.

[3] U.S. Environmental Protection Agency, "Report to congress on combined sewer overflows to the great lakes basin", U.S. EPA, Washington, DC, Tech. Rep., 2016.

[4] L. He et al., "Spatiotemporal graph convolutional network using sparse monitoring data for accurate water-level reconstruction in urban drainage systems", *Journal of Hydrology*, vol. 652, p. 132 681, 2025.

[5] S. C. Troutman, N. Schambach, N. G. Love, and B. Kerkez, "An automated toolchain for the data-driven and dynamical modeling of combined sewer systems", *Water Research*, vol. 126, pp. 88–100, 2017.

[6] T. Rosin, *Data Analytics for Automated Near Real Time Detection of Blockages in Smart Wastewater Systems*. University of Exeter (United Kingdom), 2020.

[7] V. Edmondson, M. Cerny, M. Lim, B. Gledson, S. Lockley, and J. Woodward, "A smart sewer asset information model to enable an 'internet of things' for operational wastewater management", *Automation in Construction*, vol. 91, pp. 193–205, 2018.

[8] G. J. da Silva and A. C. Borges, "Statistical process control in the environmental monitoring of water quality and wastewaters: A review", *Water*, vol. 17, no. 9, p. 1281, 2025.

[9] N. Duque, D. Duque, A. Aguilar, and J. Saldarriaga, "Sewer network layout selection and hydraulic design using a mathematical optimization framework", *Water*, vol. 12, no. 12, p. 3337, 2020.

[10] D. Meijer, M. Van Bijnen, J. Langeveld, H. Korving, J. Post, and F. Clemens, "Identifying critical elements in sewer networks using graph-theory", *Water*, vol. 10, no. 2, p. 136, 2018.

[11] D. Jung, D. Kang, J. Liu, and K. Lansey, "Improving the rapidity of responses to pipe burst in water distribution systems: A comparison of statistical process control methods", *Journal of Hydroinformatics*, vol. 17, no. 2, pp. 307–328, 2015.

[12] A. S. M. Saudi, H. Juahir, A. Azid, and F. Azaman, "Flood risk index assessment in johor river basin", *Malaysian Journal of Analytical Sciences*, vol. 19, no. 5, pp. 991–1000, 2015.

[13] R. Ibrahim, B. Saâdia, and R. Mohamed, “Comparison between different intensities of rainfall to identify overflow points in a combined sewer system using storm water management model.”, *AIMS Environmental Science*, vol. 9, no. 5, 2022.

[14] D. C. Montgomery, *Introduction to statistical quality control*. John wiley & sons, 2020.

Towards Semantic Interoperability Between IFC and CityGML Using Semantic Web Technologies

G.R. Sirigireddy and J. Krischler
Chair of Intelligent Technical Design, Bauhaus-Universität Weimar Coudraystraße 13A, 99423 Weimar, Germany
E-mail(s): ganeshwar.reddy.sirigireddy@uni-weimar.de, judith.krischler@uni-weimar.de

Abstract: This research investigates the semantic interoperability between two prominent data models: Industry Foundation Classes (IFC), predominantly used in BIM, and City Geography Markup Language (CityGML), commonly employed in GIS. Utilizing Semantic Web technologies and domain-specific ontologies such as ifcOWL, CityGML ontology, and the Building Topology Ontology (BOT), this study proposes a systematic methodology for maintaining semantic consistency across data transformations. An ETL)-based approach is employed, converting IFC and CityGML data into RDF) representations. A practical use case involving the integration of an IFC-modeled city sewer pipe segment with a CityGML-based building model demonstrates the efficacy and limitations of semantic preservation during data integration. Initial results indicate that semantic content can be preserved, providing essential connections between detailed building information and urban-scale spatial data. This work contributes to the advancement of urban information modeling by facilitating the semantic integration of parametric building models with city-level spatial data.

Keywords: Building information modeling (BIM), Geographic Information Systems (GIS), Extract-Transform-Load (ETL), Resource Description Framework (RDF), Semantic Web Technologies

DOI: 10.18154/RWTH-CONV-254897. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

The integration of BIM and GIS has gained importance in managing construction, infrastructure, and urban planning workflows. BIM provides structured, object-based representations of buildings [1], while GIS supports spatial analysis and data management at broader geographic scales [2]. However, differences in semantic models, scales, and data structures create challenges for direct interoperability between the two systems [3].

Two standards commonly associated with these domains are the IFC and the CityGML. IFC is widely used in the AEC sector to describe detailed geometric and semantic properties of building elements [4]. CityGML provides a framework for modeling and exchanging urban features in 3D, supporting multiple levels of detail for different application needs [5]. The need to combine these arises when detailed building-scale information must be analyzed alongside city-scale infrastructure data. For example, an urban planning or municipal engineering office may need to check whether newly planned

sewer pipes conflict with existing building foundations or to coordinate utility networks such as gas, water, or power lines with surrounding structures. In infrastructure projects, a transportation authority may need to evaluate how a planned bridge interacts with surrounding buildings, underground utilities, and road networks, ensuring compliance with clearance, safety, and maintenance requirements. Such analyses require both datasets to be represented in a shared, semantically consistent model that can be queried comprehensively. One of the main challenges in aligning IFC and CityGML is preserving semantic meaning during data exchange, as conventional methods often lead to information loss [6]. Semantic Web technologies, such as RDF and domain ontologies, enable the representation and linking of heterogeneous datasets while supporting querying and reasoning [7].

This research addresses semantic interoperability between IFC and CityGML through an ontology-driven approach using ifcOWL, the CityGML ontology, and BOT. An ETL process transforms data into RDF format, guided by these domain-specific ontologies to ensure semantic preservation and meaningful integration.

The approach is demonstrated through a case study integrating an IFC-modeled city sewer pipe segment with a CityGML-based building model. Results reveal both the potential and the challenges of preserving semantics in cross-domain integration, validating the proposed ETL-based RDF framework for linking infrastructure element and building data.

2 Background

Building on the foundational goal of semantic interoperability between BIM and GIS, ontology driven methods provide a promising framework to represent, integrate, and query heterogeneous datasets while preserving their semantic richness. Ontologies, which are formal models of domain concepts and their interrelationships, are central to this approach. They provide a shared, machine-readable vocabulary to resolve semantic ambiguities and enable automated reasoning. Complementing these, the RDF establishes a universal data model that encodes information as subject predicate object triples using Uniform Resource Identifiers (URIs), thereby decoupling data from original source schemas and enabling integration.

Most of the integration processes typically begin with data instantiation, wherein diverse data sources, including BIM models and GIS features, are converted into RDF triples to form a unified semantic knowledge graph [8]. For BIM, IFC is translated into ontological representations such as ifcOWL, which capture detailed semantic properties of building elements in compliance with the IFC schema [9]. Similarly, the BOT abstracts fundamental structural concepts like spaces and zones, offering consistent semantic constructs that bridge various BIM datasets [8]. This step ensures that building data is represented in a form conducive to semantic linking and inference.

In parallel, GIS datasets such as CityGML and Shapefiles are mapped onto GeoSPARQL, an OGC standard ontology designed to represent geographic features and their spatial relationships semantically [8], [9]. Although these mappings incorporate spatial context, the focus remains on maintaining semantic coherence by classifying urban elements meaningfully, such as designating pedestrian pathways or land use zones [8]. This complementary semantic encoding of BIM and GIS data establishes the groundwork for integrated urban knowledge.

Further, to reconcile semantic heterogeneity across these domains, ontology mapping employs the Semantic Web Rule Language (SWRL) to create rules that define equivalences, conditional relationships, and property mappings [8], [9]. For example, SWRL rules can align an IFC building concept with its counterpart in BOT or GeoSPARQL, facilitating consistent interpretation across datasets. Moreover, declaring BOT classes as subclasses of GeoSPARQL features enables property inheritance, thereby strengthening semantic connections between building and geographic domains [8].

At the instance level, semantic enrichment continues through the use of GeoSPARQL's topological functions, such as spatial intersections or containment, which automatically infer spatial relationships based on existing geometry references [8]. These inferred relations are appended to the knowledge graph as new semantic assertions, effectively linking indoor and outdoor spatial contexts. Additionally, the `owl:sameAs` property resolves co reference by linking instances representing the same real-world entities across disparate datasets [8].

Reasoning engines such as Apache Jena and Protégé apply these SWRL mappings and ontological definitions to derive new knowledge and implicit relationships that are not explicitly stated in the source data [9], [10]. Ensuring compatibility with reasoning mechanisms sometimes requires redefining certain properties; for instance, converting annotation properties to datatype properties to facilitate effective inference [8].

Together, these stages culminate in the creation of an integrated, semantically rich RDF knowledge base stored in triplestores like Jena TDB or Ontotext GraphDB. This knowledge base supports complex querying and retrieval through SPARQL, empowering urban management and infrastructure applications with a consistent semantic foundation [9], [10].

3 Methodology

This study presents a structured methodology for the semantic integration of heterogeneous building and urban data models, specifically those based on IFC and CityGML. The integration framework is designed around an ontology-driven ETL pipeline that converts and aligns both datasets into a unified RDF graph representation. The overall process, depicted in Figure 1, comprises three major phases: *file processing*, *semantic conversion*, and *graph-based integration*. Each phase contributes to the transformation of domain-specific datasets into semantically aligned, queryable knowledge structures.

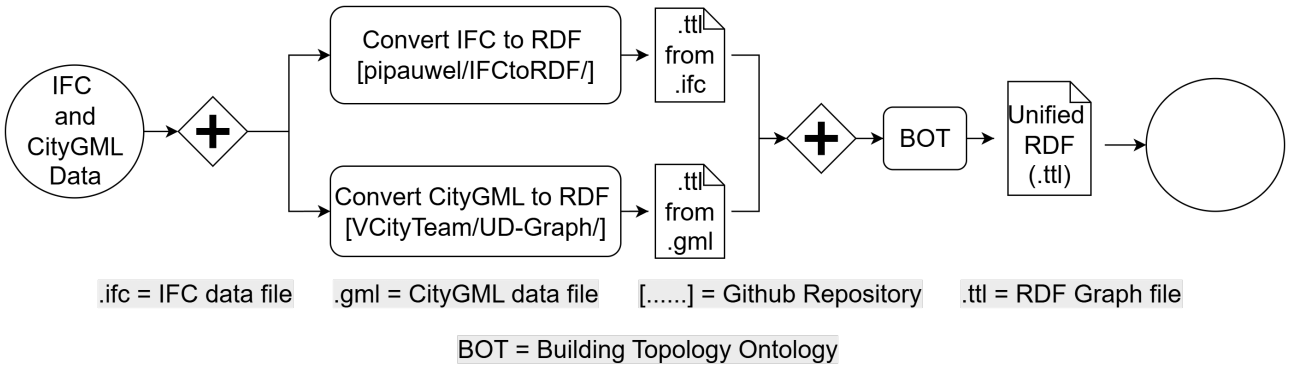


Figure 1: General architecture of the IFC-CityGML integration workflow.

3.1 File Processing and Input Validation

The workflow begins with the upload of two primary data sources: an IFC model capturing civil infrastructure components (e.g., sewer pipe segments) and a CityGML model containing above-ground building representations at Level of Detail 3 (LOD3). While the system was implemented and tested primarily with LoD3 datasets, because URIs and mapping rules were defined for this level, the same approach is applicable to LoD2. Minor adjustments to the RDF templates and mapping rules would allow LoD2 representations to be integrated without changing the overall framework. Upon upload, the system identifies the format and version of each file to ensure compatibility with supported specifications (e.g., IFC4x3, CityGML 2.0). Version metadata is parsed to route files through appropriate converters. Validation routines check for syntactic correctness, ontology consistency, and missing or malformed metadata. Files that fail validation are excluded, and detailed error reports are provided to facilitate correction.

3.2 Semantic Conversion into RDF

In the second phase, both validated datasets are independently transformed into RDF graphs. Each transformation process is aligned with relevant Semantic Web ontologies to ensure the preservation of semantic structures and relationships.

IFC to RDF Conversion: The IFC source is parsed into RDF using an IFC-specific transformation mechanism based on the `ifcOWL` ontology. This conversion preserves object hierarchies, geometrical parameters, and parametric attributes defined in the original IFC schema. Geometry-specific information and semantic descriptors are mapped to ontology terms referred to as A, through schema alignment and rule-based transformation logic.

CityGML to RDF Conversion: The CityGML model is converted to RDF using a process that applies both the `CityGML` ontology and `GeoSPARQL` vocabulary. This ensures that spatial constructs, such as surfaces and building footprints, are semantically represented using well-defined RDF triples. Thematic elements (e.g., `Door`, `RoofSurface`, `BuildingPart`) are mapped to their RDF class equivalents using JSON-based mapping rules referred to as B. For example, `<bldg:Door>` is consistently aligned to `building#Door`:

```
"{http://www.opengis.net/citygml/building/2.0}Door":
"https://dataset-dl.liris.cnrs.fr/rdf-owl-urban-data-ontologies/
Ontologies/CityGML/2.0/building#Door"
```

This ensures that semantically similar entities across the dataset use a uniform RDF vocabulary.

Each conversion results in a semantically enriched RDF graph that preserves the domain-specific semantics of the respective input models. These graphs serve as the input to the subsequent integration phase.

3.3 Integration and Semantic Alignment

The final phase of the methodology is a structured six-stage RDF integration process designed to generate a semantically interoperable knowledge graph. The integration logic is ontology-driven and

guided by a rule-based configuration that defines class mappings, property translations, spatial linking strategies, and preservation constraints.

Stage 1: All RDF graphs, including the reference ontology, are merged into a single working graph. During merging, all non-resolvable or local Internationalized Resource Identifiers (IRI) (e.g., `file:///`) are replaced with stable, ontology-aligned IRIs derived from class-level namespaces. This operation ensures long-term IRI consistency and supports linkability across external data sources.

Stage 2: Each instance in the merged graph is evaluated against declarative mapping configurations referred to as A and B to replace source-specific classes and attributes with ontology-compliant equivalents. The system applies comprehensive cross-domain alignments defined in JSON rules that map infrastructure elements such as `IfcPipeSegment` to BOT-based classes like `bot:UtilityNetwork`, and CityGML buildings to `bot:Building` with property preservation strategies. When the target ontology lacks a corresponding term, a placeholder OWL concept is generated, incorporating the missing key term from the source data as a proxy URI. This mitigates the semantic gap, albeit with inherent limitations.

A: IFC → BOT Mapping:

```
{
  "ifc:IfcPipeSegment": {
    "bot:UtilityNetwork": {
      "subClassOf": true,.....
  }
}
```

B: CityGML → BOT Mapping:

```
{
  "bldg:Building": {
    "bot:Building": {
      "equivalentClass": true,.....
  }
}
```

Stage 3: Semantic preservation rules are applied to maintain critical attributes and establish spatial relationships throughout the integration process. While Section 3.1 validates the structural correctness of the raw input files (syntax, schema compliance, and presence of basic metadata), Stage 3 ensures that all properties required for meaningful semantic integration are present in the RDF graph. It should be noted that Stage 3 relies on the input files passing Section 3.1 checks; any missing or inconsistent information in the source files can propagate into Stage 3, potentially causing incomplete semantic links or integration errors. Attributes such as `nominalDiameter`, `wallThickness`, and `measuredHeight` are used as examples; the system can be configured to validate any essential set of properties required for the analysis. Spatial relationships are established through utility validation rules that define buffer zones, depth constraints, and connectivity requirements between pipes and buildings. When pipe geometry falls within buffered regions around buildings, semantic links are created using the `servicesBuilding` predicate. Missing semantic properties are logged for review to ensure comprehensive semantic integrity.

Stage 4: The RDF graph is validated using SHACL shapes that define structural expectations and semantic constraints. Although the RDF graphs are generated beforehand, Stage 4 ensures that the integration process has not introduced inconsistencies, missing properties, or misaligned relationships. SHACL validation can fail in cases such as incomplete property propagation, unexpected null values, or violations of project-specific constraints that are not enforced during earlier stages. Violations are classified into three levels `Violation`, `Warning`, and `Info` to facilitate informed debugging and ensure compliance with the target ontology.

Stage 5: In the final step, the integrated RDF graph is cleaned, normalized, and serialized into Turtle format. A summary of integration metrics, such as triple counts, class/property replacements, and validation outcomes, is generated to support reproducibility and subsequent analysis.

This multiphase methodology ensures that heterogeneous spatial data from BIM and GIS domains are transformed into a unified semantic framework that maintains domain fidelity and supports integrative applications. Its modular design supports extensibility to additional ontologies and data types, offering a reusable foundation for smart city data integration scenarios.

4 Implementation

This section details the technical realization of the proposed integration framework using a representative use case. The implementation connects a city sewer pipe segment modeled in IFC with a residential building from a CityGML dataset. Building upon the ontology-driven ETL pipeline described in Section 3, the focus here is on the practical aspects of RDF transformation, rule-based alignment, and data serialization as applied to the selected models. The alignment of concepts is achieved by mapping both IFC and CityGML classes to a shared ontology layer based on the BOT and its extended modules. These mappings form the semantic backbone for reasoning across datasets. Mappings play the key schema-level alignments, covering both topological entities and descriptive attributes. For example, `IfcPipeSegment` is mapped to `bot-elements:SewerPipe`, while `data_measuredHeight` from CityGML is aligned to `bot-elements:height`. In addition to schema-level mappings, rule-based alignment strategies were implemented to establish dynamic semantic relationships between spatial entities. The system creates RDF graphs that preserve semantic properties from the original models.

The CityGML building maintains its semantics of geometry and EPSG:25832 coordinate reference system. The IFC pipe segment preserves its property network, including geometric parameters and material specifications. The visualization displays the interconnected structure of the converted IFC data. One key rule establishes a connection between sewer pipes and nearby buildings based on spatial proximity. This rule states that if a pipe instance lies within 3 meters of a building, a relationship is instantiated using the property `bot:servicesBuilding`. The final RDF output is organized using named graphs for modularity and traceability. Metadata about entity origin, transformation steps, and version identifiers are preserved, allowing validation and provenance tracking. Semantic queries are supported via SPARQL endpoints, enabling cross-domain data retrieval, such as querying which building is served by which underground utility or identifying connected infrastructure based on semantic class types. This implementation provides a scalable foundation for linking heterogeneous spatial data sources at the semantic level. It allows interoperability without altering the data's structural and semantic integrity. However, the current implementation supports one-way transformation from IFC and CityGML into RDF; direct reconstruction back into the original IFC or CityGML formats is not implemented. While the RDF graphs preserve the necessary schema-level classes and properties, reverse conversion would require an additional rule set to serialize RDF triples back into the native syntactic structures of the source standards.

5 Discussion

The RDF output is compatible with triple stores and graph databases, enabling BIM and GIS data queries via SPARQL. Schema-level alignment supports class and property linking across heterogeneous datasets. The pipeline is transferable to other geographic or infrastructure contexts with minimal rule file changes, facilitating interoperability in urban information systems.

However, the framework has certain limitations. The mapping logic depends on predefined schema correspondences. If source models use non-standard property names or lack metadata, alignment accuracy is reduced. For example, when IFC models omit Pset attributes or CityGML datasets lack consistent address structures, the matching process fails or requires fallback strategies. In such cases, manual rule revision becomes necessary. Moreover, RDF conversion is currently a static process: any change in the source BIM or GIS datasets requires re-running the transformation to preserve consistency, which can be resource-intensive and error-prone. Incremental update strategies or change-detection mechanisms could mitigate this by reprocessing only affected elements.

Another limitation concerns spatial metadata, which affects accurate linking of objects across datasets. When coordinate reference systems or positional anchors are missing or inconsistent, spatial linking strategies such as buffer-based matching are invalidated. Although the integration avoids direct geometry transformation, it still depends on harmonized reference frames to compute spatial relations. Disparate coordinate systems place objects in unaligned spatial reference frames, so their semantic spatial relations cannot be established without coordinate transformation. As full CRS harmonization was not performed, differing reference systems may result in incomplete or inconsistent semantic links in the graph.

In addition to spatial constraints, performance issues arise with large datasets. Processing large datasets is resource-intensive. RDF conversion from nested IFC structures generates many triples and high memory usage, while post-conversion queries slow down with long property chains or reasoning. Without indexing or triple store optimization, execution time rises sharply with model size.

In its current form, the BOT ontology does not natively support domain concepts such as utility flow direction, network connectivity, or pipe-specific properties like gradient or service depth. To represent such attributes, project-specific extensions were introduced. These are not part of any external standard and may not be interoperable with other semantic tools unless formally published and adopted. This restricts portability unless the extensions are clearly documented and shared in a compatible form. Furthermore, while the current rule set demonstrates interoperability for a representative case, it does not cover the full attribute range of IFC and CityGML. Broader applicability would require expanding the mapping rules or incorporating adaptive techniques to handle diverse model structures.

6 Conclusion and Future Work

This study introduces a semantic integration framework linking BIM and GIS datasets via ontology-aligned RDF representations. An IFC-modeled sewer pipe segment and a CityGML-based residential building were converted to RDF using ifcOWL and CityGML2OWL. Schema-level alignment with BOT and its extensions ensured interoperability, while rule-based mappings enabled semantic links between

entities without geometric fusion. The resulting knowledge graph supports cross-domain queries and traceable metadata for urban infrastructure analysis.

The case shows how static BIM and GIS datasets can be integrated through declarative rules and modular ontology mappings. Transformation logic is externalized, enabling reuse across domains through schema-level correspondence, though without validation at all levels.

The framework relies on explicit metadata and consistent property naming; deviations from standards require rule adjustments. Large datasets introduce performance constraints in RDF generation and querying. BOT's generalized structure lacks native concepts for features like utility connectivity, flow direction, or temporal states, requiring extensions for full coverage.

Planned improvements include integrating ontologies such as SAREF, FIWARE, or CityGML 3.0 ADEs to expand utility and network modeling; implementing SHACL and SPARQL-based validation for data completeness and structural integrity; and incorporating live or sensor data streams to support temporal queries and dynamic infrastructure monitoring in RDF environments.

References

[1] C. Eastman, P. Teicholz, R. Sacks, and K. Liston, *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors*. John Wiley & Sons.

[2] P. A. Longley, M. F. Goodchild, D. W. Maguire, and D. W. Rhind, *Geographic Information Systems and Science*. John Wiley & Sons, 2005.

[3] S. Herle, R. Becker, R. Wollenberg, et al., "Gim and bim", *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, vol. 88, no. 1, pp. 33–42, 2020.

[4] International Organization for Standardization, *ISO 16739-1:2018: IFC for data sharing in the construction and facility management industries*, 2018.

[5] Open Geospatial Consortium, *OGC CityGML 3.0 Conceptual Model Standard*, 2021.

[6] Y. Ji, Y. Wang, Y. Wei, et al., "An ontology-based rule mapping approach for integrating ifc and citygml", *Transactions in GIS*, vol. 28, no. 3, pp. 675–696, 2024.

[7] T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic web", *Scientific American*, vol. 284, no. 5, pp. 28–37, 2001.

[8] J. Dao, S. T. Ng, and C. Y. Kwok, "Interlinking bim and gis data for a semantic pedestrian network and applications in high-density cities", DOI: https://doi.org/10.1016/j.dibe.2024.100367

[9] J. Shi, Z. Pan, L. Jiang, and X. Zhai, "An ontology-based methodology to establish city information model of digital twin city by merging bim, gis and iot", DOI: https://doi.org/10.1016/j.aei.2023.102114

[10] S. Howell, Y. Rezgui, and T. Beach, "Integrating building and urban semantics to empower smart water solutions", DOI: https://doi.org/10.1016/j.autcon.2017.02.004

BIM- and GIS-data linkage proposals using Explainable AI (XAI) framework

G. Nikolaev[1] and J. Krischler[1]

[1]Chair of Intelligent Technical Design, Bauhaus-Universität Weimar Coudraystraße 13A, 99423 Weimar, Germany

E-mail(s): georgy.nikolaev@uni-weimar.de, judith.krischler@uni-weimar.de

Abstract: This paper presents a methodology for using locally deployed Large Language Models (LLMs) to generate meaningful, contextual links between heterogeneous data, specifically CityGML and IFC. The methodology consists of two phases: an analytical phase, where key linkage elements such as coordinate reference systems, semantic attributes, and geometric alignments are identified, and an experimental phase, where LLMs (Llama 3-8B, DeepSeek R1-8B, Gemma 3-4B) analyze the data and propose integration strategies. Beyond generating links, the approach includes an explainable AI (XAI) layer that formally documents the decision-making process using XML-based decision trees. The expected results demonstrate that the proposed approach effectively simplifies BIM and GIS data integration while ensuring transparency in the AI-driven decision-making process. By utilizing an XAI layer, the methodology addresses the black-box problem of LLM-generated content, enabling both natural language interaction and structured reverse engineering via XML-based decision trees. This ensures that engineers can analyze and interpret complex data with greater confidence.

Keywords: Semantic Web & Linked Building Data, AI & Machine Learning, BIM-GIS Integration, Large Language Models (LLMs), Semantic Linking, Explainable AI (XAI), Data Interoperability

DOI: 10.18154/RWTH-CONV-254916. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Cross-domain information continuity between Building Information Modeling (BIM) and Geographic Information Systems (GIS) is a crucial requirement for the construction industry, as it allows engineers to enhance decision-making, as data from both domains can be analyzed in context. However, there is a challenge in integrating BIM and GIS data due to the heterogeneity of their structures. Existing traditional approaches assume manual alignment and, for that reason, are labor-intensive and at risk of losing semantics. In order to address these issues, the paper proposes a methodology that uses locally deployed Large Language Models (LLMs) for generating semantic links between BIM (IFC) and GIS (CityGML), supported by an Explainable AI (XAI) layer to ensure transparency of decisions made by Artificial Intelligence (AI). The following sections will discuss the importance and challenges of BIM-GIS integration, explore existing approaches, introduce the proposed methodology, look at the

implementation details; additionally, analyze the effectiveness of the proposed approach, discuss the results, and suggest future work directions.

2 Related Work

2.1 BIM-GIS Integration

The integration of BIM and GIS has been widely explored, particularly for its value in urban planning and infrastructure management. A roadmap by buildingSMART International and the Open Geospatial Consortium emphasizes that integrating detailed BIM data with geospatial GIS contexts is essential for improved decision-making and sustainable development practices. The roadmap promotes a linked-data approach using open standards to support data integrity and BIM-GIS interoperability [1]. However, the heterogeneous nature of the GIS and BIM data makes the integration challenging. Liu et al. reviewed the integration challenges and categorized them into data, process, and application levels, clearly outlining barriers for future research [2].

2.2 Data Linking Methodologies

Data linking between heterogeneous data sources, especially IFC and CityGML or CityJSON data models, is central for BIM-GIS integration. Beck et al. identified four main approaches: conversion, extension, merging, and linking, highlighting linking as best for data independence and keeping information up to date [3]. Şenol and Gökgöz developed an automated algorithm that converts IFC to CityJSON with minimal semantic loss; however, their method focuses on conversion rather than semantic linking [4]. Lam et al. proposed an ontology-driven method that transforms IFC and CityGML into RDF graphs, using semantic web technologies to improve interoperability and 3D visualization [5]. Krischler et al. applied the ICDD standard to link BIM and GIS data in a unified framework, supporting geospatial queries and structured data aggregation [6].

2.3 Applications of Large Language Models in Data Processing

The different applications of LLMs have recently gained attention for their potential to automate complex data-related tasks. LLMs are now actively explored in engineering and construction fields for data processing. Saka et al. explored the application of GPT models in the construction lifecycle. The study identified use cases in material selection, information retrieval, scheduling, and design assistance. Although promising, the authors noted challenges such as data privacy, interoperability, and the cost of fine-tuning LLMs [7]. A recent study by Taboada et al. proposed a hybrid framework that combines vector embeddings and LLMs within a prioritized search strategy. Their retrieve-identify-prompt pipeline efficiently finds semantic correspondences without requiring model fine-tuning or complete schema alignment [8].

2.4 Explainability in AI (XAI) for Engineering and Data Integration

XAI refers to machine learning techniques that produce more interpretable models while maintaining high performance and that enable users to understand, trust, and manage AI systems effectively [9]. XAI is becoming increasingly popular in the engineering field because it helps engineers build greater trust in AI-driven decisions. Geyer et al. proposed a deep learning method for energy-efficient building design. Their approach follows systems engineering ideas and makes the model easier to interpret by using meaningful engineering values at component interfaces [10]. Additionally, Gharagoz et al.

applied a machine learning model using Extreme Gradient Boosting (XGBoost) and XAI methods to assess building performance during earthquakes. Their approach helped explain how the model works and connects the AI results with traditional structural engineering methods [11].

Previous studies have explored BIM-GIS integration through schema conversion [4], semantic ontologies [5], and container-based linking [6], these approaches face limitations due to rigid mappings and semantic loss. LLMs have shown potential in engineering data processing [7] and ontology alignment [8], yet their application for dynamic BIM-GIS linking remains underexplored. Although XAI has been used to interpret engineering models [10], [11], its integration with LLM-based linking strategies has not been addressed. This research fills this gap by combining locally deployed LLMs with XAI to propose meaningful and transparent BIM-GIS data linking.

3 Methodology

This chapter describes the methodological framework for achieving semantic linking between BIM (IFC) and GIS (CityGML) datasets. The process includes: (i) providing input BIM and GIS files together with a task-specific prompt, (ii) parsing files to reduce input size when necessary, (iii) using a locally deployed LLM to extract key attributes and construct an XAI decision tree, and (iv) generating linking suggestions, also with an XAI decision tree, which are then passed to the engineer for review (see Figure 1).

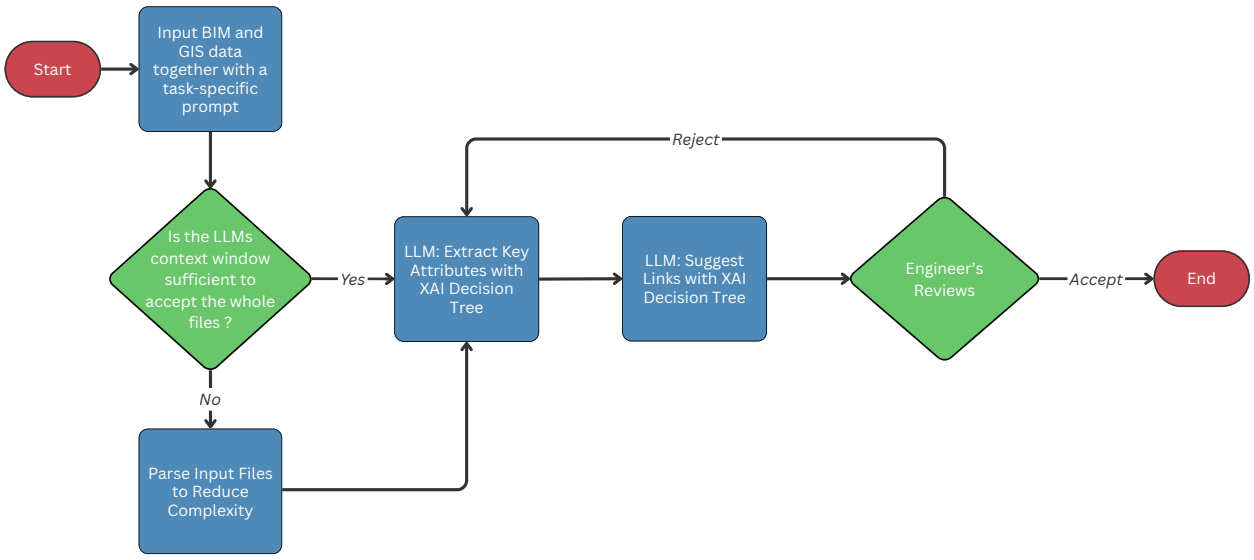


Figure 1: Overview of the methodology workflow

3.1 Linking Requirements and Data Parsing

Taking into account the heterogeneity of the BIM and GIS data discussed in the Related Work, three primary linking criteria were defined: (1) shared coordinate reference system (CRS) to ensure spatial consistency, (2) semantic attributes that describe similar entities across datasets, and (3) geometric alignment representing spatial intersections between entities.

If the LLM's context window is not sufficient to process all tokens, the IFC and CityGML files are preprocessed through a parsing step to reduce their token size. Entities likely to contain relevant information for the above criteria are extracted. For IFC, attributes such as CRS (IfcProjectedCRS), spatial hierarchies (IfcSite, IfcBuilding), coordinate transformations (IfcMapConversion), and annotations (IfcAnnotation) are extracted. For CityGML, attributes include CRS details, building footprints, characteristics (roof type, construction year, function, number of floors) and address information. This parsed content, along with a task-specific instruction prompt, is provided as input to the LLM. If the context window is sufficient, the full files with a prompt can be sent directly to the LLM.

3.2 LLM for Data Extraction and Semantic Linking with an XAI Framework

The complexity and volume of the BIM and GIS data make manual inspection time consuming and error prone. To automate the process, an LLM is used in two steps. First, the LLM helps identify and extract relevant information from IFC and CityGML files, such as coordinate reference systems, building metadata, and geometric data. Then, using this structured information, the LLM generates linking suggestions between entities in the two datasets. Model performance is improved by adjusting temperature, top-k, and top-p. For privacy of the data used, the LLM is deployed locally.

To support transparency and review by engineers, an XAI layer is included by embedding a sample output schema into the prompt. The LLM is instructed to produce structured output in XML format, following this schema.

For the data extraction step, the decision tree provides a `<Description>` of the entity being extracted, the `<ExtractionLocation>` indicating where in the original files the relevant data was found, and the `<ExtractedData>` showing the extracted values or stating `missing` if unavailable.

For the data linking step, the decision tree evaluates the extracted data on each link criteria defined in Chapter 3.1 For each candidate pair, the generated tree includes an `<Evaluation>` summary of the linking possibility, a `<Conclusion>` with one of three values: `Linked` (clear correspondence), `Inferred` (indirect or likely match), or `Not Linked` (insufficient evidence), and a `<Justification>` that explains the rationale. In the end, the tree adds a `<Suggested Strategy>` outlining the next steps for further integration.

4 Implementation and Experiments

4.1 System Setup

To start with a system setup, several software tools were chosen; their roles in the research are listed in Table 1.

A local environment was created inside a Docker container: the official *Ollama* image was pulled, the Llama 3-8B, DeepSeek R1-8B and Gemma 3-4B weights were downloaded, and the required Python packages were installed.

Note: DeepSeek R1-8B and Gemma 3-4B were added after initial experiments to compare the performance of data extraction and linking between different LLMs.

Table 1: Software used in the study

Name	Description	Role in the research
Llama 3-8B	Auto-regressive Transformer by Meta, released April 18, 2024	Information extraction; link suggestion
DeepSeek-R1-8B	Auto-regressive Transformer by DeepSeek, released January 2025	Information extraction; link suggestion; comparison
Gemma3-4B	Auto-regressive Transformer from Google's Gemma3 family, released March 12, 2025	Information extraction; link suggestion; comparison
Ollama	Open-source tool for running LLMs locally	Local execution of Llama 3-8B, DeepSeek-R1-8B, and Gemma3-4B
Python 3.10 (libraries *Langchain-Ollama*, *IfcOpenShell*, *xml.etree.ElementTree*)	Programming language with open-source libraries	Code-level interaction with LLMs; parsing of IFC and GML files
Docker	Portable, consistent runtime environment	Encapsulates all dependencies and configurations for LLMs inside an isolated container

4.2 Data Processing and Data Linking Proposals

For initial experiments with Llama 3-8B, the model's context window limit (8,192 tokens) was insufficient for the full IFC file (∼106 536 tokens), the CityGML file (∼8 478 tokens) and the task-specific prompt (∼200 tokens) together. Thus, using *IfcOpenShell* and *xml.etree.ElementTree*, the files were parsed. This reduced the IFC to 2,155 tokens and the CityGML to 3,900 tokens.

If the context window of a selected model (e.g., DeepSeek R1-8B or Gemma 3-4B) is large enough to process the full files, parsing can be skipped, and the entire file is sent to the model directly. However, since DeepSeek R1-8B and Gemma 3-4B were added after initial experiments with Llama3, to keep future model comparison consistent, the methodology was kept similar across all models. That is, the initial parsing was present in all models.

The original IFC-STEP file was first converted to IFCXML to produce a flatter structure, simplify traversal, and align file formats between BIM and GIS inputs. The parsed data are then wrapped in an explainable XML structure described in Chapter 3.2 and passed together with the prompt to the LLM for linking suggestions.

The explainable-AI component for linking takes the form of a decision tree as described in Chapter 3.2 with three main criteria, `<SharedCRS>`, `<SemanticAttributes>` and `<CoordinateAlignment>`.

4.3 Prompt Engineering

Communication with the LLM is handled through the *Langchain-Ollama* Python library by passing the task-specific instruction prompt. The Google prompt engineering paper guideline [12] is followed by combining the system, role, and one-shot prompting. The prompts are organized according to the task type:

- **Data extraction** prompts instruct the model to act as a BIM/GIS specialist and to return results as an XML decision tree. The LLM is guided to extract the relevant information from the IFC-XML and CityGML files for the criteria (see Chapter 3.1), with each step containing a description, extraction location, and extracted value.

- **Data linking** prompts instruct the model to act as a BIM/GIS integration specialist and to compare the extracted information. The model is required to output an XML decision tree containing the evaluation of linking possibilities and providing a conclusion and justification for each criterion.

The temperature is set to 0 in order to obtain deterministic answers (the model always selects the next most probable token). Parameters such as *top-*k and *top-*p were evaluated, but with the temperature fixed at 0 they no longer affect the output.

5 Results and Discussion

5.1 Data Extraction and Links-accuracy evaluation

The extracted data was manually checked and all relevant attributes were correctly extracted according to the linking criteria defined in the prompt. For the suggested links evaluation, a ground-truth standard was manually constructed. According to this standard, the correct conclusions are: `<Not Linked>` for CRS, `<Linked>` for Semantic Attributes, `<Linked>` for Geometric Alignment, and ultimately the Suggested Strategy should be `<Inferred>` with justification to add CRS to the IFC file. The results are summarized in evaluation Table 2, where accuracy is assessed based on four criteria: CRS, Semantics, Geometry, and Strategy. Additionally, an observation column was introduced to state general patterns in each model's output.

Table 2: Links-accuracy evaluation

Model	CRS	Semantics	Geometry	Strategy	Accuracy	Observations
Llama3-8B	Inferred	Not Linked	Inferred	Inferred	25%	Information incompleteness
DeepSeek-R1-8B	Inferred	Inferred	Linked	Inferred	50%	Cautious assertions
Gemma3-4B	Not Linked	Inferred	Linked	Inferred	75%	Cautious assertions

To demonstrate the structured and transparent reasoning provided by the XAI layer, Figure 2 shows the CRS evaluation node from the decision tree generated by the best-performed model Gemma3-4B. Only this model correctly concluded that there is no direct link between the files due to the lack of information in the IFC-file.

```
<DecisionTree>
    <SharedCRS>
        <Evaluation>IFC data lacks CRS information. GML uses EPSG:25832.</Evaluation>
        <Conclusion>Not Linked</Conclusion>
        <Justification>The IFC data does not specify any CRS. The GML data uses
         EPSG:25832, which is a projected coordinate system. Without a common CRS, a
         direct link is not possible. Assigning EPSG:25832 to the IFC data would be a
         necessary step for linking, but it's currently missing.</Justification>
    </SharedCRS>
    ...
</DecisionTree>
```

Figure 2: Example of an explainable decision node(CRS evaluation) from Gemma3-4B

As a result, Llama3 omitted some relevant semantic and coordinate details, leading to incomplete conclusions. DeepSeek-R1 was generally correct, but cautious in its assertions, often going for `<Inferred>` instead of stating `<Linked>`. Gemma3 aligned the closest with the ground truth, although

it also used a cautious conclusion for semantic matching. Notably, none of the models produced hallucinations, e.g. no information was made up, which was not always the case initially. This improved after applying prompt engineering practices.

5.2 Discussion

Consequently, we obtained three XML files, each enriched with a decision tree: the first two store the extracted and organized IFC data and CityGML data respectively, and the third contains the proposed links between them with justifications. All LLM models handled summarization and extraction of key information reliably, although the quality of further reasoning and link proposals was noticeably better with newer models, whose answers matched the ground-truth standard more closely. Notably, none of the models was tuned before testing, and conversational context was not stored between the prompts. The only technique applied was prompt engineering, using role, system and one-shot prompting; additionally, the prompts were kept similar across models. Moreover, the determinism of extracted data at scale is a valid concern; an alternative is to use the LLM to generate executable queries against the source data while still preserving the decision tree explainability.

Surprisingly, the model with the smallest number of parameters, Gemma3-4B, performed best, achieving the highest accuracy relative to ground truth. Performance results may suggest that Gemma3-4B's training data align better with BIM-GIS contexts, although this cannot be confirmed without access to the model's training corpora. The gap for accuracy which we observed might be filled after enhancement of the prompt engineering or fine-tuning of the models. However, so far, engineers can use the proposed framework as a starting point to validate extracted data, compare it directly with the source files, and check how each proposed link between BIM and GIS data was generated.

6 Conclusion

In this paper, the novel framework to achieve semantic linking between heterogeneous BIM (IFC) and GIS (CityGML) datasets is proposed, using locally deployed LLMs and an integrated XAI layer. The methodology uses LLMs for data extraction, structured analysis, and the generation of linking suggestions, where each step is documented with a transparent XML decision tree. The results demonstrate that LLMs are capable of extracting relevant attributes from complex engineering data and suggesting links that closely align with the human solution.

In addition, an XAI layer enables engineers to interpret AI-generated suggestions, which addresses the black-box limitations usually associated with LLMs. However, there is still room for improvement; future work could focus on automatic self-validation of results, fine-tuning the models to enhance accuracy and perform a deeper justification; also, one could experiment with more advanced prompt engineering practices and evaluate newer, more capable LLMs. The current framework can be expanded in order to analyze additional data formats containing BIM and GIS data; thus, making the application wider and more informed about the content. Hence, the discussed approach offers a promising step towards more efficient and transparent BIM-GIS data integration.

Acknowledgement

We thank Dr. Jakob Taraben and Prof. Christian Koch for their valuable advice and feedback.

References

[1] J. Mallela and A. Bhargava, *Enabling information continuity across bim-gis domains: A bsi and ogc strategic roadmap*, https://docs.ogc.org/wp/24-057.pdf, 2024.

[2] X. Liu, X. Wang, G. Wright, J. Cheng, X. Li, and R. Liu, "A state-of-the-art review on the integration of building information modeling (bim) and geographic information system (gis)", *ISPRS International Journal of Geo-Information*, vol. 6, no. 2, p. 53, 2017. DOI: 10.3390/ijgi6020053

[3] S. Beck, J. Abualdenien, I. Hijazi, A. Borrmann, and T. Kolbe, "Analyzing contextual linking of heterogeneous information models from the domains bim and uim", *ISPRS International Journal of Geo-Information*, vol. 10, no. 12, p. 807, 2021. DOI: 10.3390/ijgi10120807

[4] H. Şenol and T. Gökgöz, "Integration of building information modeling (bim) and geographic information system (gis): A new approach for ifc to cityjson conversion", *Earth Science Informatics*, vol. 17, pp. 3437–3454, 2024. DOI: 10.1007/s12145-024-01343-1

[5] P.-D. Lam, B.-H. Gu, H.-K. Lam, S.-Y. Ok, and S.-H. Lee, "Digital twin smart city: Integrating ifc and citygml with semantic graph for advanced 3d city model visualization", *Sensors*, vol. 24, no. 12, p. 3761, 2024. DOI: 10.3390/s24123761

[6] J. Krischler, P.-C. Schuler, J. Taraben, and C. Koch, "Using icdd for bim and gis integration in infrastructure", in *LDAC2024 - Linked Data in Architecture and Construction*, M. Poveda-Villalón and P. Pauwels, editors, 2024.

[7] A. Saka et al., "Gpt models in construction industry: Opportunities, limitations, and a use case validation", *Developments in the Built Environment*, vol. 17, p. 100 300, 2024. DOI: 10.1016/j.dibe.2023.100300

[8] M. Taboada, D. Martinez, M. Arideh, and R. Mosquera, "Ontology matching with large language models and prioritized depth-first search", *arXiv preprint arXiv:2501.11441*, 2025. [Online]. Available: https://arxiv.org/abs/2501.11441

[9] D. Gunning, "Explainable artificial intelligence (xai)", Defense Advanced Research Projects Agency (DARPA), Tech. Rep., 2017. [Online]. Available: https://www.darpa.mil/attachments/XAIProgramUpdate.pdf

[10] P. Geyer, M. M. Singh, and X. Chen, *Explainable ai for engineering design: A unified approach of systems engineering and component-based deep learning*, https://ssrn.com/abstract=4833100, 2024.

[11] M. M. Gharagoz, M. Noureldin, and J. Kim, "Explainable machine learning (xml) framework for seismic assessment of structures using extreme gradient boosting (xgboost)", *Engineering Structures*, vol. 327, p. 119 621, 2025. DOI: 10.1016/j.engstruct.2025.119621

[12] L. Boonstra, *Prompt engineering*, https://cloud.google.com/generative-ai/docs/learn/prompts, 2024.

Discrete Event Simulation (DES)-Based Abstraction of Construction State Machines and Robotic Simulations for Site-Wide Optimization Using BIM and Gazebo

Ankita Maurya[1]
[1]Institute of Civil Engineering, Technische Universität Berlin, Germany
E-mail(s): a.maurya@tu-berlin.de

Abstract: The construction industry is one of the most labor-intensive and least digitized and automated industries worldwide, facing low productivity, worker shortages, and safety issues. Since the last decade, automation and robotics have witnessed great development in building their capacity to enhance productivity, sustainability, and adaptability on-site. To allocate a large number of human and robotic resources, a strategy needs to be designed for an entire construction site. Such planning activities require the development of a representation of elements at a granularity level that considers detailed geometrical data, the robotic equipment placement, and the safety of all the workers.

Therefore, the paper proposes an integrated collaborative approach of computational architecture in the form of a Building Information Model (BIM), a simulation interface of Gazebo for developing a robotic construction site, and real-time Discrete Event Simulation (DES) for optimization of resources and evaluation of safe construction spaces. As a part of the research, a case study is selected where the construction of prefabricated factory halls is planned, considering the resources like robots, delivery trucks, cranes, and the on-site human fleet. Thus, the paper will contribute to an elaborate integration of optimizing robotic construction processes and logistics supply chain planning.

Keywords: DES, Robotic simulation, BIM, Automation in Construction, Gazebo.

DOI: 10.18154/RWTH-CONV-254915. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

1.1 Problem Statement

The construction industry, which is vital to Europe's economy, has been much slower to apply advanced and scientific methods for analyzing complex construction processes, most likely due to the labor shortages, skills mismatches, low productivity growth, and slow technology uptake [1]. Automating processes in the construction sector requires the application of many different robots and equipment to work in close collaboration.

Construction sites with automated equipment, robots, and human workers face high uncertainty, affecting activity durations, material delivery, machine reliability, and labor productivity [2]. To manage this complexity, strategic resource planning is essential. One effective method is creating a virtual project model using simulation. This allows decision-makers to analyze workflows, allocate resources, and manage risks. Among simulation types: System Dynamics, Agent-Based Modeling, and Discrete Event Simulation (DES), DES is widely used for addressing such uncertainties.

DES, a computational method, models a system's changing behavior and performance as a sequence of discrete events over time, offering granular insight into equipment placement, material deliveries, or human–robot interactions. However, DES alone can be limited in resource allocation or design, requiring manual updates to the model, resulting in time-consuming revisions and limiting industry adoption. Research has shown that linking DES with a central BIM database allows real-time updates and automated model changes [3]. To address uncertainty and variability in such dynamic environments, stochastic models can be incorporated, such as Monte Carlo simulations, Weibull, and logistic distributions. Monte Carlo methods allow for probabilistic risk analysis through repeated sampling, while Weibull distributions help model equipment lifespans or failure probabilities, and Logistic distributions capture decision-making under uncertainty. Including these models enhances the robustness of DES by accounting for random events and probabilistic behavior. Moreover, for real-time updates to the simulation model, it is essential to establish a connection to a central database, achievable through the integration of BIM with the overall computational architecture.

BIM is a central digital tool for designing, planning, and managing construction projects, containing detailed geometry and data. When integrated with robotic simulation, it supports automated construction. However, while BIM has advanced, less focus has been placed on linking it with robot operations and semantics. Technologies like Robotic Operating System (ROS) are enabling greater robot adoption, but strategic planning methods for full-site robotic work are still lacking. Therefore, effective planning will require simulation models that can represent various robotic tasks and allow user-controlled actions across different levels.

1.2 Research Objective

The objective of this research was, therefore, to develop and validate an integrated simulation framework that combines BIM, robotic simulation using Gazebo, and DES to support the planning, coordination, and optimization of construction site operations. The framework aimed to abstract complex construction workflows into discrete state machines, enabling dynamic evaluation of resource allocation strategies, spatial safety constraints, and construction sequencing in real-time. By applying this framework to a real-world case study involving prefabricated factory hall construction, the study aimed to demonstrate its effectiveness in improving operational efficiency and site-wide automation.

2 State of the Art - Literature Review

2.1 BIM and Robotic Systems

Recent research has increasingly focused on the simulation of robotic construction tasks. To date, much of this work has emphasized the development of strategic simulation models for factory-based prefabrication. Notable examples include the frameworks proposed by [4] and [5], which integrate

robotic production planning with established BIM standards to simulate the prefabrication of building components. Similarly, [6] developed a BIM-based simulation model for the prefabrication of wooden frames.

In contrast, relatively few studies have addressed the simulation of on-site robotic construction activities. Among those, a key area of interest lies in developing interfaces between BIM and ROS to enable task-level coordination. These studies typically address specific, narrowly scoped applications. For instance, [7] demonstrated BIM–ROS integration for simulating automated indoor painting. [8] investigated robotic masonry using a cable-driven system, while [9] proposed a simulation framework for robotic excavation.

2.2 Discrete Event Simulation in construction planning

DES models complex systems as chains of sequential states, each representing a process step. In construction, a crane lifting operation might follow: (a) fastening the load, (b) lifting it, and (c) installing it. State transitions occur after time delays or external events. The DES formalism extends this by including resource constraints, such as material availability or truck arrivals.

Applying DES in construction faces two key challenges: (a) accurately modeling construction sequences and (b) accounting for variable activity durations due to site dynamics. The first requires linking DES with BIM to reflect design-specific sequences and geometries [10]. The second involves integrating real-time data from sensors to model productivity variability [11].

Given these challenges, DES is best viewed as part of a hybrid modeling approach, effective for automated and robotic workflows, especially when integrated with BIM and simulation tools like Gazebo.

2.3 Gaps in Current Research

In summary, there is limited research exploring how robot-assisted construction work can be effectively represented and simulated. Existing studies largely focus either on modeling prefabrication processes in controlled factory settings or on highly specific robotic automation tasks. However, they rarely address how such tasks can be integrated and situated within the complex, dynamic environment of construction sites, an environment that requires stochastic modeling to capture its inherent variability. In the next section, a framework is proposed to integrate these tasks and apply the techniques to real-world situations.

3 Research Methodology

3.1 Conceptual Framework of the Study

The proposed framework integrated two main domains: data-driven BIM and cyber-physical robotic production resources, enabling realistic simulation of on-site robotic construction tasks. The integration process was organized into three key components: the BIM interface, the DES environment, and the simulation loop, as defined in Figure 1.

BIM data, exported in Industry Foundation Classes (IFC) format, served as the foundation for the simulation. The BIM Interface extracted essential parameters such as task identifiers, site coordinates, element types, and material quantities. This parsed data was processed by an IFC File Parser, which transformed it into construction-related inputs, including schedules and logistics data, that were then

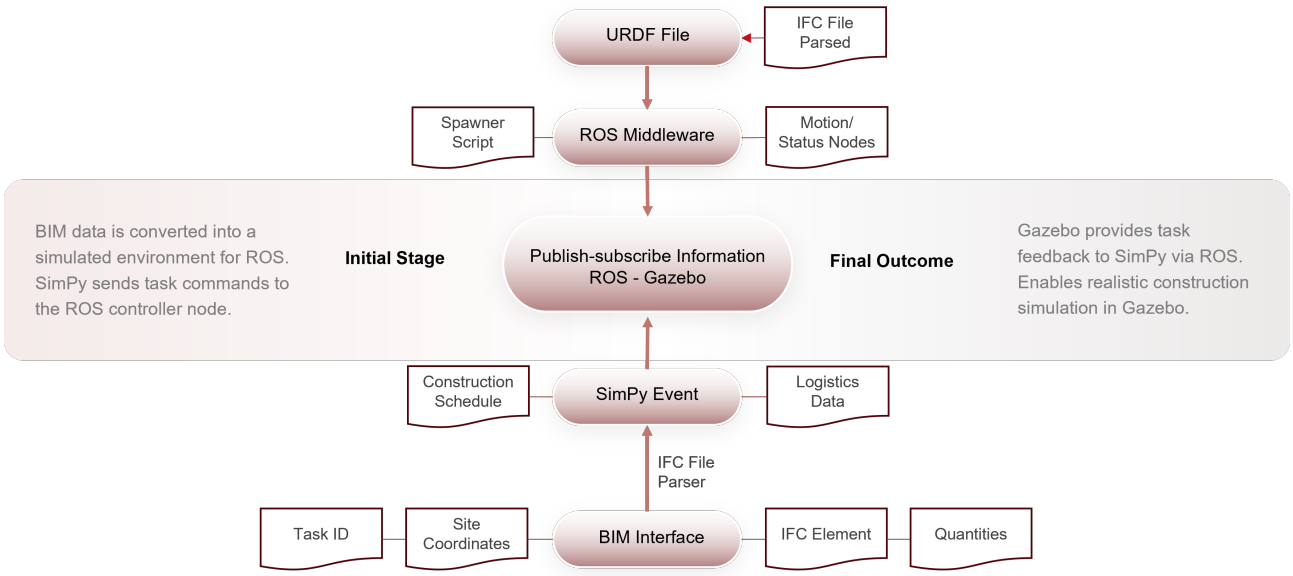


Figure 1: Data flow from BIM to DES to Gazebo

input into Python-based SimPy to generate discrete construction events. Additionally, the construction schedule was also added to the DES model, which was then implemented in the SimPy library. Within SimPy, stochastic processes and Monte Carlo sampling allow the system to estimate task durations, resource utilization, and performance variations under different site conditions.

The ROS-Gazebo middleware acts as the bridge between the DES and the simulated robotic resources. The models in the format of Unified Robot Description Format (URDF) were parsed from BIM-derived geometry and spawned into Gazebo via ROS spawner scripts. Motion and status nodes enabled continuous bidirectional communication: SimPy sends task execution commands to ROS controllers, while Gazebo provides real-time task feedback back to SimPy.

This publish–subscribe integration enabled dynamic simulation where task duration, logistics, and operational constraints are updated during runtime. The framework not only estimated productivity and efficiency for robotic operations but also allowed planners and designers to test alternative resource configurations, site layouts, and weather delay scenarios before physical placement.

3.2 Case Study Context and Data Modeling

To validate the framework, a case study from the Horizon 2020 ASHVIN project was used, which focused on the construction of a prefabricated factory hall spanning 30,000 square meters and 12 meters in height. The dataset included the two-week delivery and mounting schedule of prefabricated columns. A robotic agent was integrated into the SimPy-based DES model to perform mounting tasks alongside human workers, while a robotic tower crane was modeled for placing the columns into position.

Upon the truck's arrival at the construction site, the operational workflow followed the logic shown in Figure 2. First, parking availability was checked; if no space was available, the truck waited in a parking queue until one was freed. Once parked, the system verified the availability of critical

resources—namely, the tower crane, robots, human operators, and the designated mounting location. If any of these were occupied, the truck remains in the parking queue until all were released.

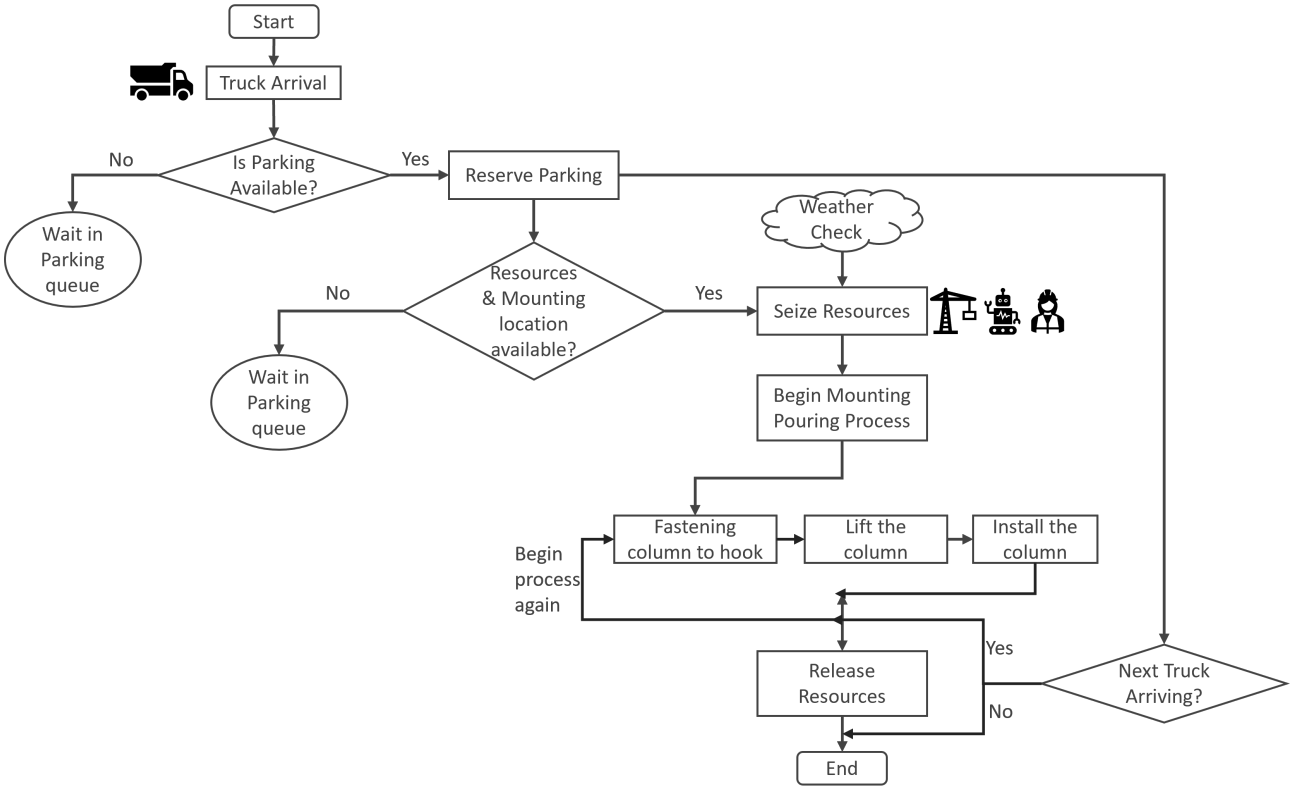


Figure 2: DES model for the crane

When parking and resources were available, a weather check was performed to ensure safe operating conditions. In the event of extreme weather, all activities were paused until conditions improved. If conditions were acceptable, the resources were seized and the mounting process began. The duration of each mounting cycle was determined using a probability distribution function (PDF), defined here as a statistical model representing the likelihood of different task durations based on historical records, sensor measurements, and stochastic analysis. Sampling from this PDF allows the simulation to capture real-world variability in operator speed, alignment precision, and minor delays, rather than assuming fixed times.

The mounting cycle consists of fastening the column to the crane hook, lifting the column, and installing it in position with the help of a robot and a human. This process repeats for each column transported by the truck. Once all columns were installed, the resources and parking slot were released, the truck departs, and the system transitioned to the next truck's arrival, repeating the sequence.

4 Implementation and Results

The framework was developed using Python 3.10 with the SimPy library for handling discrete-event scheduling, resource management, and queuing processes. Weather conditions were included as external factors influencing the simulation flow. In the model, key resources such as robots, human workers, cranes, trucks, and parking spaces were represented with capacities matching the case study

setup. The workflow was divided into three main steps: a delivery truck arriving and parking, a crane lifting and positioning a column, and a worker–robot team mounting the column in place. Each step required the relevant resources to be available, and SimPy's request/wait mechanism ensured that tasks were queued when resources were busy. Task durations were based on realistic timing data and included variation to reflect real-world uncertainty.

To evaluate system performance under uncertainty, the model was tested using a Monte Carlo approach. For each experimental setup, the simulation environment was reset and run until all columns were assembled, with the total completion time recorded at the moment the final column was mounted. To assess how different resource allocations affect productivity, multiple scenarios were created by varying the numbers of robots, workers, cranes, and trucks. Each scenario was simulated repeatedly to capture the effects of randomness, and the results were averaged to provide reliable performance estimates.

4.1 Results

Three categories of resources: robots, cranes, and trucks, were evaluated for their impact on project completion time. The results were obtained from 100 Monte Carlo runs per configuration to account for stochastic variability in delivery and mounting durations. The simulations examined the effect of varying human and robotic resource allocations on project completion times, accounting for operational constraints such as traffic and weather delays.

Robots vs Workers Allocation: The heatmap in Figure 3 illustrates the average completion times for different combinations of robots (2 to 8 units) and human workers (4 to 10 units). Across all configurations tested, average completion times ranged narrowly between 3146 and 3153 minutes, indicating that the marginal gains from increasing either resource type alone were minimal under the tested operational model. Notably, certain configurations with moderate resource levels (e.g., 6 robots and 6 workers) yielded slightly longer completion times than others with lower or higher counts, suggesting nonlinear interactions between workforce composition and task sequencing efficiency.

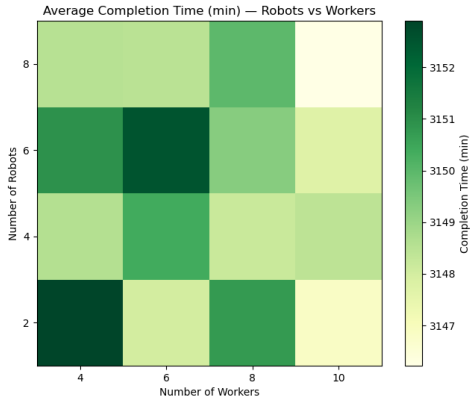


Figure 3: Average completion time

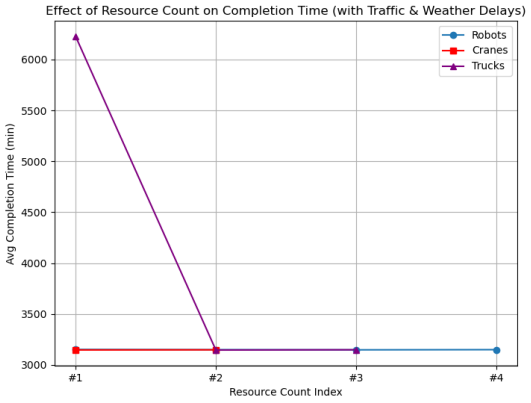


Figure 4: Effect of Resources on Completion Time

Effect of Resources on Completion Time: Figure 4 presents the influence of resource counts for robots, cranes, and trucks when traffic and weather delays were introduced. Here, the effect of truck availability was most pronounced: with a low truck count, completion times peaked at over 6200 minutes, more than doubling the baseline. Increasing truck count immediately reduced completion times to the same level as the best-performing crane and robot configurations (3110 minutes). This emphasizes that logistical constraints in material delivery can overshadow productivity gains from on-site automation alone. From a general safety perspective, the ability of the framework to simulate spatial constraints in real time allows informed placement of robots and human workers, maintaining operational safety margins even under high-density deployment.

5 Conclusion and Future Work

Although the framework included plans for integration with the Gazebo simulation platform, work is still in progress to support more realistic robotic kinematics, advanced environment sensing, and robust collision detection. While the results demonstrated clear potential for improving resource planning, several limitations remain. The current simulation environment assumed idealized equipment performance and operator reliability, which may not fully reflect the effects of mechanical failures, human error, or learning curves. Disturbance modeling was restricted to generic traffic and weather delays, whereas real-world scenarios often involved more complex factors such as variable delivery lead times, on-site congestion, and unexpected environmental hazards.

Future work will address these limitations by integrating probabilistic failure models, extending the model to capture complex logistical uncertainties, and coupling the spatial safety module with real-time perception and adaptive path-planning algorithms. Additionally, validating the framework against real-world project data will be critical to assessing its predictive accuracy and operational feasibility. These enhancements will strengthen the maturity of the framework and support its adoption as a decision-support tool for safe, efficient, and adaptive human–robot collaboration in construction.

Acknowledgments

The research presented in this paper has received funding from the Deutsche Forschungsgemeinschaft as part of Research Unit 5672, "Digital Backbone of Robotized Construction".

References

[1] J. M. Davila Delgado et al., “Robotics and automated systems in construction: Understanding industry-specific challenges for adoption”, *Journal of Building Engineering*, vol. 26, p. 100 868, 2019. DOI: https://doi.org/10.1016/j.jobe.2019.100868 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352710219300889

[2] M. Wu, J.-R. Lin, and X.-H. Zhang, “How human-robot collaboration impacts construction productivity: An agent-based multi-fidelity modeling approach”, *Advanced Engineering Informatics*, vol. 52, p. 101 589, 2022. DOI: https://doi.org/10.1016/j.aei.2022.101589 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1474034622000611

[3] W. Lu and T. Olofsson, “Building information modeling and discrete event simulation: Towards an integrated framework”, *Automation in Construction*, vol. 44, pp. 73–83, 2014. DOI: https:

//doi.org/10.1016/j.autcon.2014.04.001 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580514000879

[4] A. Zhu, P. Pauwels, and B. De Vries, "Component-based robot prefabricated construction simulation using IFC-based building information models", *Automation in Construction*, vol. 152, p. 104 899, Aug. 1, 2023. DOI: 10.1016/j.autcon.2023.104899 Accessed: Jan. 23, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580523001590

[5] W. Anane, I. Iordanova, and C. Ouellet-Plamondon, "BIM-driven computational design for robotic manufacturing in off-site construction: An integrated design-to-manufacturing (DtM) approach", *Automation in Construction*, vol. 150, p. 104 782, Jun. 1, 2023. DOI: 10.1016/j.autcon.2023.104782 Accessed: Jun. 2, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580523000420

[6] O. Wong Chong, J. Zhang, R. M. Voyles, and B.-C. Min, "BIM-based simulation of construction robotics in the assembly process of wood frames", *Automation in Construction*, vol. 137, p. 104 194, May 1, 2022. DOI: 10.1016/j.autcon.2022.104194 Accessed: Jan. 23, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092658052200067X

[7] K. Kim and M. Peavy, "BIM-based semantic building world modeling for robot task planning and execution in built environments", *Automation in Construction*, vol. 138, p. 104 247, Jun. 1, 2022. DOI: 10.1016/j.autcon.2022.104247 Accessed: Jan. 23, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580522001200

[8] T. Bruckmann and R. Boumann, "Simulation and optimization of automated masonry construction using cable robots", *Advanced Engineering Informatics*, vol. 50, p. 101 388, Oct. 1, 2021. DOI: 10.1016/j.aei.2021.101388 Accessed: Jun. 2, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1474034621001282

[9] D. Guan, N. Yang, J. Lai, M.-F. F. Siu, X. Jing, and C.-K. Lau, "Kinematic modeling and constraint analysis for robotic excavator operations in piling construction", *Automation in Construction*, vol. 126, p. 103 666, Jun. 1, 2021. DOI: 10.1016/j.autcon.2021.103666 Accessed: Jun. 2, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580521001175

[10] S. Spieckermann, "Intelligent BIM-based construction scheduling using discrete event simulation", *Winter Simulation Conference*, Dec. 9, 2012. DOI: 10.5555/2429759.2429836 Accessed: Jun. 2, 2025. [Online]. Available: https://www.academia.edu/115993518/Intelligent_BIM_based_construction_scheduling_using_discrete_event_simulation

[11] A. S. Rao et al., "Real-time monitoring of construction sites: Sensors, methods, and applications", *Automation in Construction*, vol. 136, p. 104 099, Apr. 1, 2022. DOI: 10.1016/j.autcon.2021.104099 Accessed: Jun. 2, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580521005501

Unsupervised Semantic Segmentation of Cracks in Concrete Specimens

Amir Sadrnia[1], Mohamed S. H. Alabassy[1] and Andrea Osburg[1]
[1]Bauhaus-Universität Weimar, F. A. Finger Institut für Baustoffkunde, Professur Bauchemie und Polymere Werkstoffe
E-mail(s): mohamed.said.helmy.alabassy@uni-weimar.de

Abstract: Reliable and automated detection of cracks is increasingly used for quantitatively assessing the condition of internal damage to concrete subject to cyclic freezing and thawing. Numerous methods already exist to detect cracks, but semantic segmentation has demonstrated superior results in comparison to other conventional methods utilizing heuristic image processing techniques. However, training such networks requires large human-annotated datasets. To eliminate the need for labor-intensive pixel-annotation of images, unsupervised semantic segmentation offers an alternative solution. In this research, two state-of-the-art frameworks, Self-supervised Transformer with Energy-based Graph Optimization (STEGO) and Eigen Aggregation Learning (EAGLE), were chosen to be trained on the μCT and DeepCrack datasets due to their success on benchmarks. The quantitative results of STEGO trained on μCT are promising, but the qualitative results show that it failed to detect them, whereas EAGLE with nearly identical metrics could partially detect cracks, when class imbalances are not addressed. Based on the qualitative results on DeepCrack, both models detected cracks in this dataset, however, EAGLE performed worse than STEGO in metrics.

Keywords: Crack segmentation, Unsupervised learning segmentation, Contrastive learning, Vision transformer

DOI: 10.18154/RWTH-CONV-254881. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Cracks may occur when water gets absorbed into the voids of a heterogeneous multi-phase composite porous material, like concrete, then freezes causing pressure forces inside, as the volume of freezing water expands exerting expansive force beyond its tensile strength. To efficiently assess freezing durability and serviceability of concrete, a significant number of approaches, varying from traditional and manual techniques to image-based methods, were employed to achieve crack detection.

Although previous works, particularly in image-based methods, have achieved significant success in crack identification and detection by development of Deep Learning (DL) methods, automated crack detection is receiving close review and of current interest. Deep learning algorithms can automatically learn meaningful features from the images and are capable of handling large-scale data and complex

scenes, where such approaches ranging from supervised semantic segmentation to unsupervised semantic segmentation were used to tackle the automated crack detection tasks.

Supervised semantic segmentation approaches are reliable methods for crack detection tasks, however, these methods require manually labelled data in order to train the deep learning models. Unsupervised semantic segmentation models, on the other hand, can segment meaningful features without the need for annotated masks. These approaches include reduced reliance on extensive labeled datasets and human expertise, adaptability to diverse environments, the potential for real-time applications in infrastructure maintenance and enables scalable and cost-effective crack detection, especially for large infrastructures.

In this research the potential of unsupervised techniques to detect cracks in concrete specimens is explored. State-of-the-art models Self-supervised Transformer with Energy-based Graph Optimization (STEGO) [1] and Eigen Aggregation Learning (EAGLE) [2] were chosen for their significant performances on benchmarks [3]. These models have demonstrated promising results in various applications showcasing their ability to handle complex data.

2 Methodology

2.1 Dataset

One of the datasets used for this study was obtained at Bauhaus-Universität Weimar, Chair of Construction Chemistry and Polymer Materials. The dataset consists of μCT-scans of a selection of cylindrical concrete specimens, which were exposed to cyclic freezing and thawing according to the standard German procedures for the Capillary Suction, Internal Damage and Freeze Thaw (CIF) test, then scanned using the “nanotom m research | edition” micro and nano computed tomography system. The system, housed in a compact and radiation-shielded cabinet, is equipped with a high-power nanofocus X-ray tube and an internally cooled detector for extremely high magnification and resolution.

A total number of 3115 grayscale slices of 3D image stack and their corresponding masks were generated. Each mask consists of two classes for segmentation, crack and non-crack. Approximately 2% of the dataset belongs to cracks, whereas 98% are assigned as non-crack pixels. The segmentation scheme was adopted to enable detailed analysis of the internal structures and damage processes within the concrete samples.

The second dataset used in this research, DeepCrack [4], consists of 537 RGB images of concrete surfaces showing a varying levels of crack severity. The images include differences in exposure, brightness, shadow, texture and crack characteristics, and resemble to a large extent the morphological features cracks captured in computed tomographic scans. It reflects the real-world conditions, where cracks may occur in various environments. The images also vary in resolution and quality to ensure that the model encounters a wide range of data during training. This dataset also contains two classes, crack and non-crack, and almost 3% of the the images are labeled as cracks and the rest are non-crack regions.

2.2 Preprocessing

Preprocessing the dataset involves manipulating, cleaning, filtering, and encoding the data in order to feed the input to the network, and ensuring that it is free of errors, such as missing or NaN values. The

μCT dataset was originally grayscale so they were converted to RGB values to ensure the compatibility with Self-distillation with no labels (DINO) [5] backbones. Furthermore, all images were set to a fixed resolution of 224$\times$224 pixels for a balanced memory usage and computational efficiency as well as ensuring the compatibility with DINO backbones. Moreover, for consistent intensity ranges and the reduction of the effect of lighting variations can be achieved through normalization. Normalization takes the mean and standard deviation of the pixel values and normalize them to a fixed range of [0,1] or [-1,1].

Annotation or consistent label mapping is essential for aligning with the model's predictions for meaningful evaluation. In this research, the masks of the datasets are not used for training, however, they are important for an accurate evaluation, therefore, label mapping is a vital preprocessing technique. CT reconstruction artefacts filled or inconsistent images are considered outliers and may cause errors during training, the presence of noisy data could contribute to over-fitting and severely impact the model's performance in learning meaningful features. The μCT dataset was reviewed and the noisy data were manually removed. After this step, the total number images were downsampled from 3154 Pixels to 2512. A reduction of almost 20% made the dataset cleaner and prepared it for the next preprocessing steps.

The datasets in this study both consist of images and annotation masks. It should be noted that the labels of the datasets are never used during the training process; however they are solely used for evaluating the performance of the model [1]. The μCT dataset was divided into two parts; training and validation datasets. A 70:30 data-split ratio (i.e., 70% of the dataset is used for training and the other 30% for validation purposes) was taken into account, while a ratio of 80:20 was considered for the DeepCrack dataset, as the whole dataset contains significantly fewer images in compared to the μCT dataset. For μCT, 1759 images and their corresponding labels belong to training dataset, whereas 753 images and their corresponding annotation masks belong to validation dataset. For DeepCrack, the numbers are 300 and 237 for the training and validation datasets respectively. Data augmentation is used to artificially generate more varying images from the existing data. Zooming-in, jittering and cropping out images can simulate different data acquisition conditions and geometric perspectives and reduce bias. In STEGO and EAGLE data augmentation is carried out by five-crop technique, which divides an image into 5 different crops, top-left corner, top-right corner, bottom-left corner, bottom-right corner, center crop with desired resolution.

2.3 Models

STEGO is a Self-Supervised Learning (SSL) model, which uses Vision Transformers (ViT) instead of Convolutional Neural Network (CNN) to detect patterns in images without relying on manually annotated data. EAGLE takes a different approach by clustering similar features in images by using eigenvector analysis. Both models have demonstrated solid performance on benchmark segmentation tasks, showcasing their potential for real-world application, such as crack detection. However, the repository of EAGLE was developed based on that of STEGO.

Both models leverage contrastive learning. STEGO learns how to semantically cluster a group of similar pieces, whereas EAGLE aggregates features by identifying shared patterns across different images, which can be compared with recognizing dominant eigenvectors in a dataset. For the

purpose of this study, a refactored version of STEGO is used for training on the respective datasets. Frey et al. [6] refactored the STEGO project to achieve a clean and more organized code. Their implementation is more efficient in usage of Random Access Memory (RAM) and is therefore more time efficient for training purposes of the model locally. Training of the EAGLE and STEGO models on DeepCrack and μCT datasets was conducted on a laptop with NVIDIA GeForce RTX 4080 (12 GB) Graphics Processing Unit (GPU), Intel Core i9-13980HX Computing Processing Unit (CPU), and 64 GB DDR5-5600 (2x 32 GB Dual-Channel) of RAM.

2.4 Evaluation Methods

To evaluate the results of a semantic segmentation process, numerous metrics can be employed. In STEGO and EAGLE, the performance of the trained models are evaluated through Accuracy and mIoU. Accuracy (Acc) is calculated by the number of correct predictions over the whole scene. It is calculated using the formula $Acc = \frac{TP+TN}{TP+TN+FP+FN}$, while Intersection over Union in terms of semantic segmentation calculates the overlap of the ground truth and the prediction (True Positive) over the union of the sample sets. The formula for Mean Intersection over Union (mIoU) is $IoU = \frac{TP}{TP+FP+FN}$, where TP, TN, FP and FN are the number of True Positive, True Negative, False Positive and False Negative values respectively.

Both STEGO and EAGLE are evaluated using two primary methods, namely the Linear Probe Evaluation and Clustering Evaluation. The linear probe assesses the quality of the learned features via a supervised task by training a linear classifier on top of a frozen feature extractor using Cross-Entropy loss, with ground-truth labels used only for evaluation [1]. In contrast, clustering evaluation measures the alignment between unsupervised predictions and ground-truth labels using the Hungarian algorithm to ensure label permutation invariance, to evaluate the model's ability to group similar regions meaningfully [1].

3 Experiments and Results

The STEGO model with the architecture *"DINO ViT_base_8"* was trained on the μCT and DeepCrack datasets with 500 number of steps, batch size 32 and 0 and 5 extra clusters, additional learned representations help to improve segmentation performance, for μCT and DeepCrack respectively. After hyperparameter tuning, it was determined that the architecture, batch size and number of extra clusters are the key parameters influencing STEGO's training. The base version of DINO with 8 patches showed the best performance related to mIoU and accuracy among the tested backbones. Increasing the batch size to 32 could significantly improve the model's performance. The value of extra clusters can influence the quantitative and qualitative results of the training. Training STEGO on the μCT dataset with zero extra clusters yielded the best performance for mIoU.

EAGLE with the architecture *"DINO ViT_base_8"* was also trained on μCT and DeepCrack using a batch size of 32, with 600 and 500 maximum training steps, and 5 and 0 extra clusters, respectively. Similar to the findings regarding the sensitivity and importance of hyperparameters in STEGO, EAGLE exhibited comparable behavior with respect to its own hyperparameter settings. It is noteworthy that the parameters b_{knn}, b_{rand}, b_{self}, γ_{knn}, γ_{rand}, and γ_{self}, negative inter-cluster alignment, positive inter-cluster relationship, positive intra-cluster alignments, negative inter-shift, positive inter-shift and positive

intra-shift respectively, demonstrated significant importance in the hyperparameter optimization of STEGO. Among these, the parameters b_{knn}, b_{rand}, and b_{self} were particularly crucial for optimizing EAGLE's performance. The quantitative and qualitative results of training STEGO and EAGLE on μCT and DeepCrack are shown in Table 1, Figure 1, Figure 2 and Figure 3 respectively.

Table 1: Quantitaive results of Training STEGO and EAGLE on μCT and DeepCrack datasets respectively.

Model	**Dataset**	**Unsupervised**		**Linear Probe**	
		Accuracy (%)	mIoU (%)	Accuracy (%)	mIoU (%)
STEGO	μCT	92.11	46.07	97.40	48.70
STEGO	cropped-μCT	97.81	49.89	99.85	50.02
STEGO	DeepCrack	98.96	49.53	99.68	49.90
EAGLE	μCT	99.96	45.33	99.23	48.89
EAGLE	cropped-μCT	99.98	45.76	99.97	50.11
EAGLE	DeepCrack	68.49	34.25	99.83	49.91

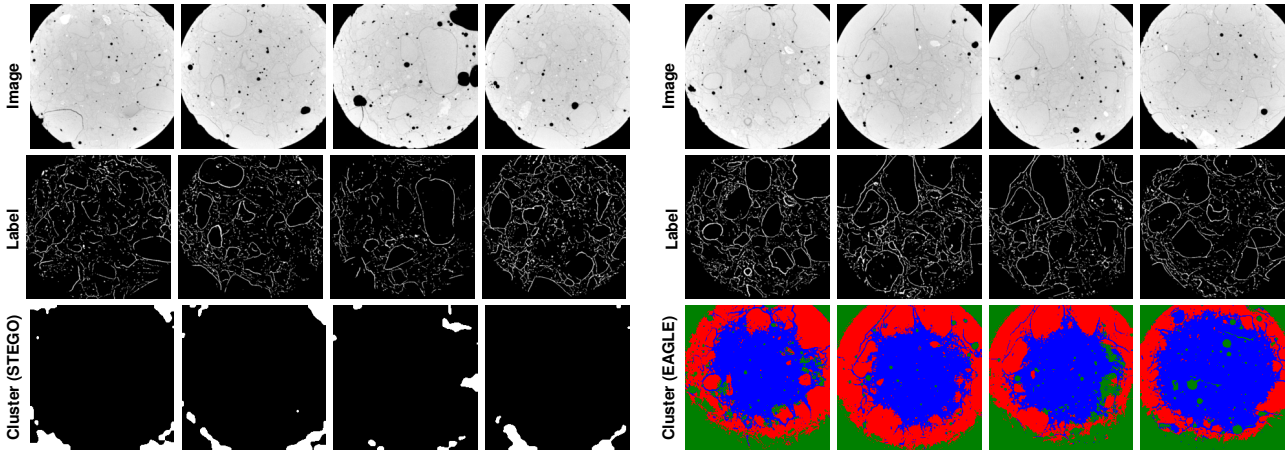


Figure 1: Qualitative results of the models on μCT dataset. On the right STEGO and on the left EAGLE. The blue color for results of EAGLE on Cluster depicts the 'Crack' class, whereas red demonstrates the anything on the concrete except for cracks. Green does not indicate a third class but it was used to distinguish the background from the sample.

4 Discussion

4.1 Comparison of the Results

The evaluation of the experiments shows that STEGO could provide a reasonable quantitative results for both datasets, but the qualitative results show that STEGO failed to detect any cracks in the μCT dataset, whereas it detected cracks in the DeepCrack dataset. EAGLE produced nearly identical quantitative results for the μCT dataset but detected partially some cracks in the images. EAGLE showed a weaker quantitative performance for DeepCrack, however its qualitative results were promising. It was found that the models' metrics alone do not fully reflect the segmentation quality attained due to the disproportionality of class weights, but through visualization of the results and how

the model segments the respective classes, the performance of the approach can be interpreted more accurately.

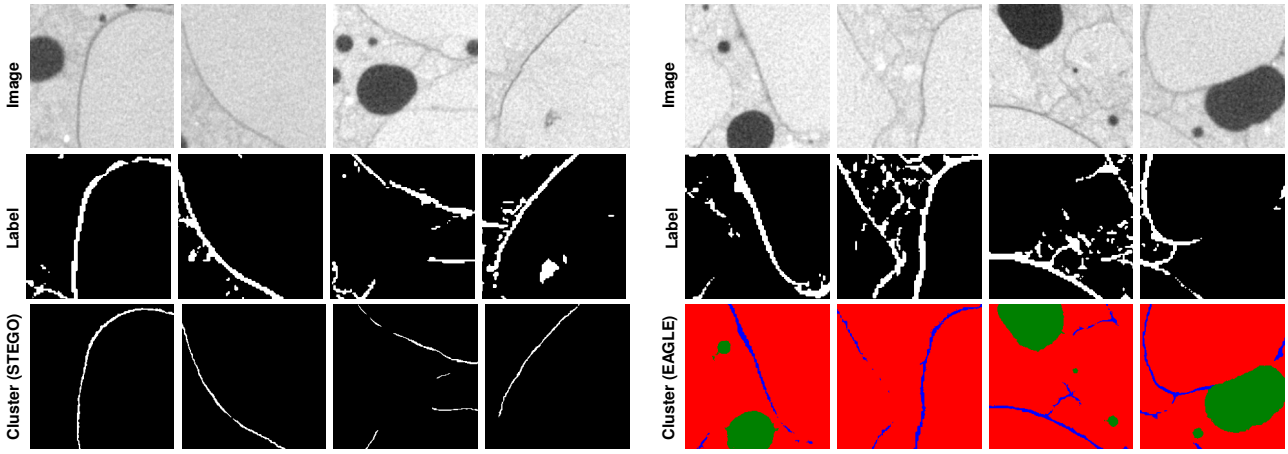


Figure 2: Qualitative results of the models on cropped μCT dataset. On the right STEGO and on the left EAGLE. The blue color for results of EAGLE on Cluster depicts the 'Crack' class.

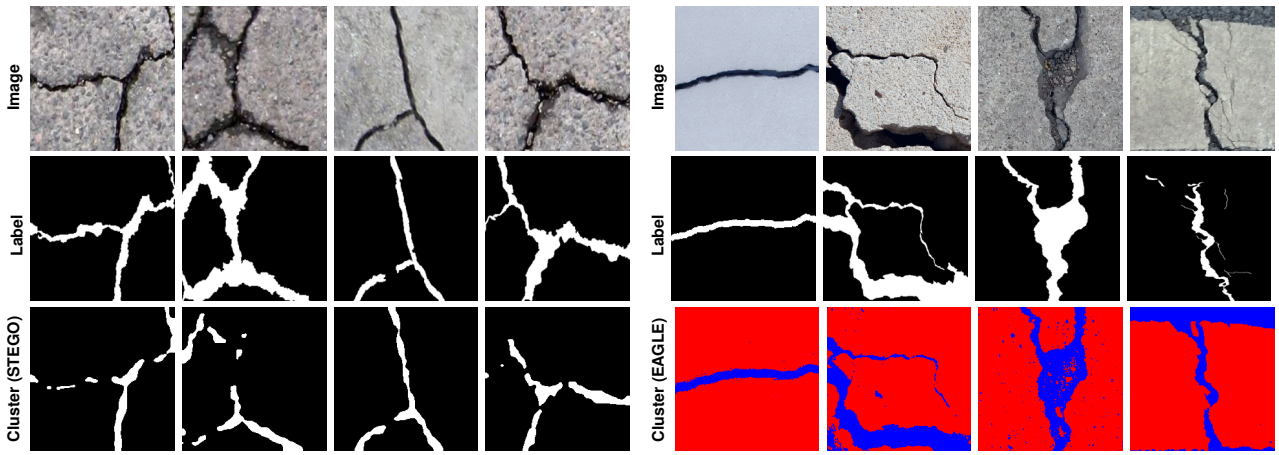


Figure 3: Qualitative results of the models STEGO and EAGLE on DeepCrack dataset.

4.2 Challenges and Countermeasures

4.2.1 Data-Level Approaches

Despite successful application of unsupervised semantic segmentation using STEGO and EAGLE, several challenges were encountered that affected the performance of the models. The greatest challenge is the class imbalance. The class imbalance encourages the model to become biased and leads to overfitting. The second challenge mostly regards the μCT dataset and its one-channel (grayscale) structure. Most of the existing models are designated to be trained on RGB datasets. However, several strategies were implemented to attack the aforementioned problems.

The FiveCrop augmentation was applied to the DeepCrack dataset to increase diversity by cropping each image into five parts. However, for μCT images, which contain cracks, concrete, and background but still follow a binary classification (crack vs. non-crack) it was avoided to prevent segmentation challenges. Instead, standard cropping was used to reduce background interference and enhance crack visibility. This improved both models' mIoU by approximately 3% each.

Manually removing the images with a lesser presence of cracks was employed for DeepCrack. This approach would allow the dataset to magnify the crack class. This approach could improve the mIoU STEGO and EAGLE by approximately 1% and 0.9%. It should be noted that the conventional augmentation techniques, such as rotation, translation and jittering was applied during the training.

Grayscale structure of μCT was also problematic, as most of the semantic segmentation models expect RGB images. However, duplicating grayscale channels to generate fake RGB images was taken into account. This approach could only make the dataset appropriate for being fed to the models. Furthermore, transfer learning could not be effective because pre-trained models are trained on RGB datasets. Moreover, traditional color-based augmentations such as hue and saturation could not be applied to grayscale images.

4.2.2 Algorithm-Level Approaches

Switching between DINO architectures was explored to improve model performance. While DINOV2 is theoretically more advanced, it did not consistently achieve high metrics in this study. The best results were scored by *"ViT_base"* with 8 patches. Adding extra clusters helped STEGO on DeepCrack (+4% mIoU) and EAGLE on μCT (+2% mIoU), but the same strategy failed when applied conversely. This highlights dataset-dependent model behavior. Koenig et al. [7] suggest that a lower-dimension embedding space, D_{STEGO}, focuses on global pattern and generates coarse segmentations, whereas higher values keep more granular details but causes noise and therefore a less effective segmentation. The target value 90 for *"dim"* showed the best performance for all four scenarios.

It is noteworthy to mention that the difference between STEGO and EAGLE in detecting cracks, even partially, in μCT dataset lies in their approaches to capture meaningful data. EAGLE employs EiCue (Eigen Cues) as a spectral clustering technique. EiCue helps EAGLE segment sharp, fine structures rather than just grouping similar features. STEGO's feature correspondence approach works better to distinguish larger semantic regions but may fail dealing with thin cracks. For reference, the UP-CrackNet method [8] achieved an IoU of 58.738 trained on DeepCrack, and 61.073 trained on Crack500, demonstrating strong performance among unsupervised approaches. These IoU values are higher than those obtained in our experiments, although the qualitative results from UP-CrackNet are less clear in capturing fine crack details.

5 Conclusion

In this research, the unsupervised models STEGO and EAGLE were trained on μCT and DeepCrack datasets for binary crack segmentation, showing that while STEGO achieved better quantitative results and EAGLE yielded more promising qualitative outputs, visual inspection revealed that segmentation quality, aside from metrics values, better reflects model performance. The difference between STEGO and EAGLE performance lies in their techniques to capture cracks. EAGLE utilises an EiCue (i.e., Eigen Cue) technique, which avails the model to extract sharper and finer details, whereas STEGO relies on feature correspondence approach, which is more suitable for distinguishing larger semantic regions. Furthermore, the qualitative differences in model performance, particularly the failure of STEGO to detect cracks in μCT, indicates that class imbalance play the central role. Numerous methods such as data-level approaches as well as algorithm level approaches were employed. Minor

improvements were observed after applying the techniques discussed earlier, but these techniques could not fully solve the problem.

Future research could explore alternative methods to overcome class imbalance. Moreover, more robust, less noisy and class-balanced datasets might help improve crack detection. It should be noted that unsupervised approaches are still in their early stages and the current results motivate further improvements in the architectures and algorithms. It is still too soon to draw conclusions, whether such approaches are appropriate or the definitive solution for crack detection. Nevertheless, the completion of this research represents a significant step toward leveraging existing models for crack detection. In conclusion, this research demonstrates the potential of unsupervised techniques for crack detection, opening new avenues for further exploration in this field.

Data Availability and Acknowledgements

Trained checkpoint files of the trained models available on GitHub. A Master's Thesis pertaining further detailed results of this paper is available via this URL. This research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project number 432767097. μCT Dataset used is part of the aforementioned project.

References

[1] M. Hamilton, Z. Zhang, B. Hariharan, N. Snavely, and W. Freeman, *Unsupervised semantic segmentation by distilling feature correspondences*, arXiv preprint, arXiv:2203.08414, 2022.

[2] C. Kim, W. Han, D. Ju, and S. Hwang, *Eagle: Eigen aggregation learning for object-centric unsupervised semantic segmentation*, arXiv preprint, arXiv:2403.01482, 2024.

[3] Papers with Code, *Unsupervised semantic segmentation*, Available at: https://paperswithcode.com/task/unsupervised-semantic-segmentation.

[4] Y. Liu, J. Yao, X. Lu, R. Xie, and L. Li, "Deepcrack: A deep hierarchical feature learning architecture for crack segmentation", *Neurocomputing*, vol. 338, pp. 139–153, 2019. DOI: 10.1016/j.neucom.2019.01.036

[5] Yannic Kilcher, *Dino: Emerging properties in self-supervised vision transformers (facebook ai research explained)*, 2021. [Online]. Available: https://youtu.be/h3ij3F3cPIk

[6] J. Frey, M. Mattamala, N. Chebrolu, C. Cadena, M. Fallon, and M. Hutter, *Fast traversability estimation for wild visual navigation*, 2023. arXiv: 2305.08510 [cs.RO]. [Online]. Available: https://arxiv.org/abs/2305.08510

[7] A. Koenig, M. Schambach, and J. Otterbach, *Uncovering the inner workings of stego for safe unsupervised semantic segmentation*, 2023. arXiv: 2304.07314 [cs.CV]. [Online]. Available: https://arxiv.org/abs/2304.07314

[8] N. Ma, R. Fan, and L. Xie, *Up-cracknet: Unsupervised pixel-wise road crack detection via adversarial image restoration*, 2024. arXiv: 2401.15647 [cs.CV]. [Online]. Available: https://arxiv.org/abs/2401.15647

Hygrothermal Performance of Clinker Bricks with Perforations: A Comparative Simulation Study

Andrea Thiele[1], Daniel Kastner[1], Oliver Abele[1] and Jérôme Frisch[1]
[1]Institute of Energy Efficiency and Sustainable Building (E3D), RWTH Aachen University, Aachen, Germany
E-mail(s): andrea.thiele@rwth-aachen.de, daniel.kastner@rwth-aachen.de, abele@e3d.rwth-aachen.de, frisch@e3d.rwth-aachen.de

Abstract: Modern façade systems that actively contribute to a building's energy performance are gaining importance. A notable example is a wall-integrated pipe system used for solar energy collection. While thermal aspects of such systems have been widely studied, hygrothermal behavior—including moisture-related risks such as condensation, frost damage, or algae growth—has received comparatively little attention. This study addresses this gap by exploring how to best represent perforated clinker bricks containing an embedded pipe system in a two-dimensional hygrothermal simulation. Therefore, two modeling approaches are compared: vertical sections that include mortar joints and horizontal sections that include the brick's solid material layers. The pipe system remains inactive to focus on the complex structure of clinker, mortar, and perforations. Simulations using WUFI 2D are validated with experimental data. Due to missing hygrothermal material data for clinker, standard brick data is used and adjusted for both model types. The results show that neither the vertical nor the horizontal section provides a universally superior representation. Instead, each approach has specific advantages and drawbacks, the vertical model has a better representation of the absorption process at the moisture saturated weathered side of the clinker and the horizontal model shows a better distribution.

Keywords: hygrothermal simulation, moisture, perforated clinker brick, WUFI, facade systems

DOI: 10.18154/RWTH-CONV-254914. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

While traditional building façades mostly play a passive role of separating the indoor from the outdoor climate, new façade systems in which the façade plays an active role within the building's energy system are gaining in relevance. These systems often work by embedding pipe systems in the walls that allow heat flows to be directed within the wall.

Most of the existing literature on pipe embedding systems focuses on purely thermal simulations to assess the system's performance and potential. Especially systems that aim to direct heat gains from a wall into a heat pump, effectively cooling down the wall, and therefore also interfere with the building part's moisture balance. This might increase the risk of condensation or slowing down the

drying process, which in turn might cause moisture induced damage such as frost, salt-crystallization spalling or algae growth on the wall. Therefore, a hygrothermal investigation of these types of systems is needed. The literature on embedded piping focuses primarily on thermal simulations, with hygrothermal aspects receiving limited attention. Vogdt and Helbach stand out by using a 1D hygrothermal model in WUFI Pro, treating embedded pipes as heating elements [1]. Further, Ibrahim et al. combine numerical and experimental approaches to investigate the real-world behavior of systems that transfer solar gains from the south facade to the north [2].

In this paper, a clinker brick façade that embeds pipes beneath the outer surface is investigated in the scope of the ReFaTEk research project [3] with the aim to analyze moisture dynamics and conduct a risk assessment. Over the course of a day, the clinker wall is heated by the solar radiation and reaches temperatures higher than the surrounding air. This heat is absorbed by brine flowing through the pipes and used to supply a heat pump. Brick walls are by nature very heterogeneous due to the perforations and the surrounding solid material layers as well as the mortar in between bricks. This paper investigates how this heterogeneous geometry can be translated into a 2D hygrothermal simulation model using the software WUFI 2D by comparing multiple different cross sections of the wall. Since in-situ monitoring of a wall's moisture content is not directly possible, the humidity and temperature of the air measured inside the perforations within the clinker are used to evaluate the model performance.

2 Experimental Setup

For the verification and validation of the models, a preliminary test rig of the ReFaTEk project is used to investigate measurement concepts for moisture measurement where the humidity in the perforations of the clinker bricks are measured and the surface is exposed to simulated rain events.

A clinker containing sensory equipment (sensor clinker) was developed and tested to continuously measure the humidity and temperature inside the clinker. The used measurement data was gathered by sensors installed in perforations from below by routing the cables through the mortar. Additionally, a simple 3d-printed grid cover was made and installed to protect against falling mortar. The inserted Bosch BME280 humidity sensor is built in a breakout board from BerryBase. The sensor can measure humidity with an accuracy of ± 3 % RH from 20 to 80 % RH at 25 °C, as well as temperatures from 0 °C to 65 °C with an accuracy of ± 0,5 °C. All sensors used were batch calibrated. The test rig is five clinkers wide and tall as shown in Figure 1. The black sensor clinker in the center of the wall is the one used for all further experimental data. In addition to the sensors in the front and rear perforations, there is a BME280 ambient sensor that measures the temperature and humidity of the air around the test rig. To simulate a rain event, a measured and controlled amount of water was sprayed evenly onto the front of the clinker test rig.

After taking a few measurements, the test rig was carefully dismantled to examine the performance of the sensor modules. It could be seen that the mortar had penetrated the perforations further than expected. In the third picture of Figure 1 the mortar is pressed about 12 mm into the perforation with a sensor from below. Therefore, the sensor is still closer to the mortar despite the successful protection

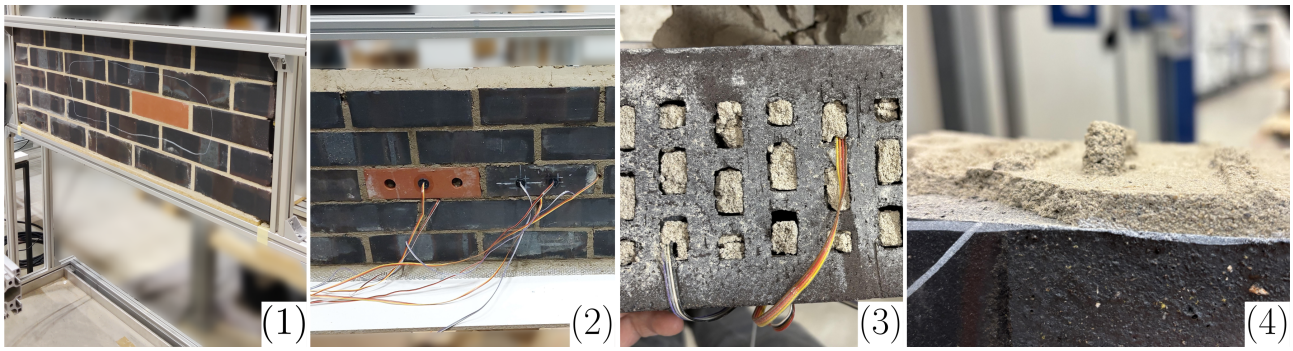


Figure 1: Test wall with sensors in the circled black and red brick from the front (1) and back (2). On the right (3 & 4) the penetration of mortar deep into the perforations, even from below, can be seen.

from above due to the grid cover. Thus, the humidity measurements in the perforations as well as the moisture transport could be strongly influenced by the mortar.

3 Simulation

The simulations are carried out using WUFI 2D, a two-dimensional hygrothermal simulation software based on the coupled heat and moisture transport model developed by Künzel [4]. Since no validated clinker brick material data is available in the WUFI database, the material data used in this study is based on the properties of an extruded clay brick. To more accurately reflect the behavior of the tested clinker, several adjustments were made to the hygrothermal material functions. In particular, the moisture storage function was modified on the basis of measured data from a commissioned Dynamic Vapor Sorption analysis by the Institute of Building Materials Research (ibac) at RWTH Aachen University in order to better reflect the sorption properties observed in the experiment. Additionally, both the redistribution of moisture within the material and the capillary suction behavior were adapted based on the experimentally determined maximum moisture content that the clinker brick was able to absorb in the experiments of the ibac. For the extruded brick, the initial maximum water content was 370 kg/m³, whereas the clinker brick showed a significantly lower maximum value of 72 kg/m³. The adaption was simply calculated by using the devisor of these values and applying it to the curves.

To simulate driving rain loads on the façade, the recorded weather data from the experiment is applied, including temperature, relative humidity and wind-driven rain. A targeted 15-minute rain impulse is introduced at a defined point in the simulation to replicate short, intense splash water events. In addition, the rain was extended by a further 3 minutes of minimal rain to represent the water adhering to the façade, because WUFI 2D assumes an instant dry surface without an active rain event. To represent the vapour diffusion resistance of the dense surface of clinker bricks, a small parametric simulation study was conducted. Although this is not presented in this paper, it was found that applying an sd-value of 0.1 m to the outer surface improved the results. This value was then applied to all subsequent models. As the relative humidity and temperature have not reached an equilibrium between the measurements inside the perforations and outside the bricks, starting conditions for the simulation are interpolated linearly across the brick layers between the sensors.

To investigate moisture transport, a comparative study of different two-dimensional simulation approaches is conducted. The study is structured in two phases: First, several vertical and horizontal section models are analyzed and compared within their respective groups. Second, additional mortar is introduced into the perforations of the best performing models of each group to investigate the impact of the mortar entering the perforations found during deconstruction of the test rig. The different model configurations are explained in the following section and visualized in Figure 2. The vertical sections include the horizontal mortar layers above and below the bricks but neglect the material layers between perforations. Four vertical model configurations are considered:

- Vertical Model 1 assumes perfectly aligned large perforations stacked on top of each other, separated only by mortar joints.
- Vertical Model 2 introduces a more realistic arrangement with alternating large and small perforations at the front face, representing the offset pattern of actual brickwork.
- Vertical Model 3 includes a solid material layer beneath a large perforation, representing the fact that perforations will not always be stacked on top of each other.
- Vertical Model 4 accounts for excess mortar filling the air space next to the pipe, reflecting situations where mortar falls into perforations or enters the perforations from the bottom during installation.

The horizontal sections, by contrast, incorporate the brick's solid material layer structure, which are not captured in vertical models. Three configurations are examined:

- Horizontal Model 1 slices through the middle of the brick, cutting directly through the material layers between perforations. This allows for analyzing horizontal moisture flow through the brick's inner structure.
- Horizontal Model 2 places the section just below the embedded pipe, within the continuous mortar layer. In this view, the pipe itself is not intersected, so its material properties do not block moisture flow toward the rear of the brick.
- Horizontal Model 3 adds excess mortar filling within the perforations to Horizontal Model 2.

By comparing these models first within their respective orientations and then across orientations, the study aims to clarify the role of specific geometric features—such as perforation alignment, mortar joints, and material layer—in shaping moisture transport. Since the aim of this study is mainly to compare different modelling approaches, a qualitative analysis of the resulting simulation data is deemed sufficient. Additionally, the root mean squared error (RMSE) is calculated to provide further insight and a quantitative comparison. The RMSE is calculated separately for the water absorption and the drying phase. The beginning of the drying phase has been defined as the last timestep where the maximum humidity is measured during the experiment. This leads to slightly different starts of the drying phases for the front and rear perforations.

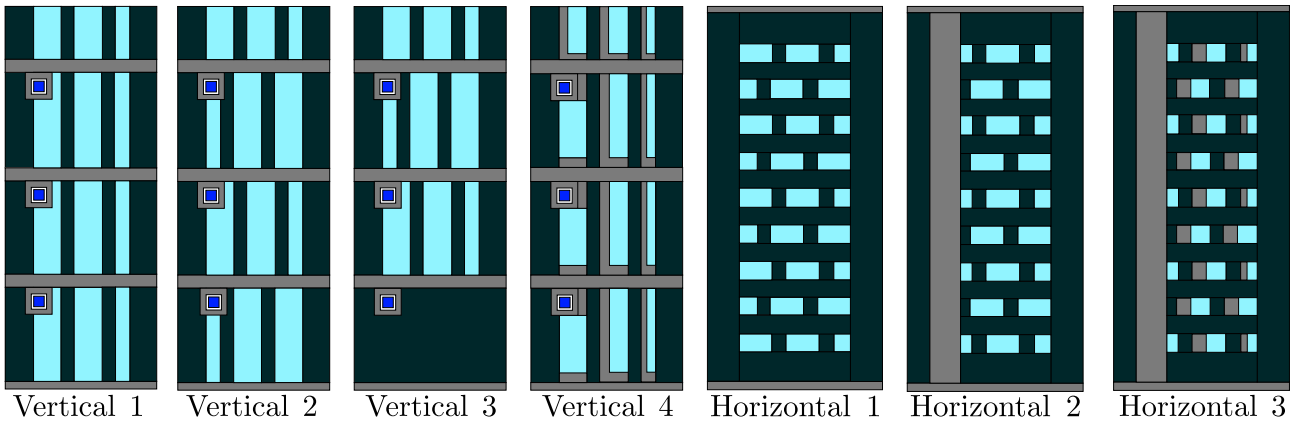


Figure 2: Illustration of the different vertical and horizontal models

4 Results

To evaluate all models the simulation data is compared to the experimental data of the black sensor clinker qualitative and quantitative with RMSE values. All models displayed a very similar behaviour of the temperature curves with RMSE values of 0,48 °C to 0,51 °C for models over the complete simulation time. Therefore, the focus of the results will be on the relative humidity.

4.1 Comparison of vertical modelling approaches

Within the vertical section simulations, different stacking configurations of the bricks were analyzed. In all simulation variants, the relative humidity in the front perforation increased more rapidly than in the experiment, and the values in the rear perforation were consistently higher than measured. These deviations suggest an overestimation of both the absorption rate and the transport of moisture toward the rear of the brick in the simulations. While the graphs in Figure 3 show almost identical profiles during the humidity absorption process, they start to differ during the drying process. Vertical model 2 (with alternating perforations) starts drying first for both the front and rear perforation and has the steepest drying curves out of the three models. Vertical model 1 (aligned perforations) and 3 (solid clinker brick in the bottom row) show very similar curves with model 1's drying process starting the latest, thus leading to the highest humidity values. The RMSE during the drying process of model 1 and 3 of 8.36 % and 8.58 % indicate a slightly better performance of vertical model 1 for the front perforation while vertical model 3 shows better performance for the rear perforation (RMSE Model 1: 6.47 %, Model 3: 5.86 %). Since the differences in performance are relatively small, the simpler model (Vertical Model 1) is used for the rest of this study.

Both temperature and moisture simulations show a rapid convergence towards the ambient conditions. This effect can be seen both at the start and towards the end of the simulation time, where all simulations converge to environmental conditions while the measured values lie significantly above this value. This rapid convergence remains a limitation and highlights an open issue regarding model calibration and initial condition definition.

4.2 Comparison of horizontal modelling approaches

Two different horizontally sectioned models were simulated and compared. The humidity curves for the models are shown in Figure 4. A direct comparison between the two horizontal models

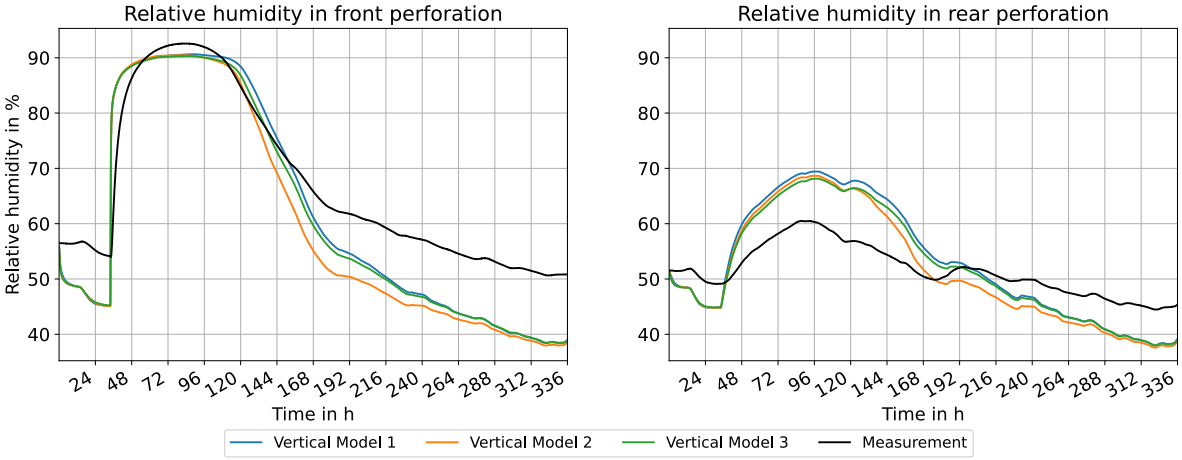


Figure 3: Relative humidity in front and rear perforation of Vertical Models

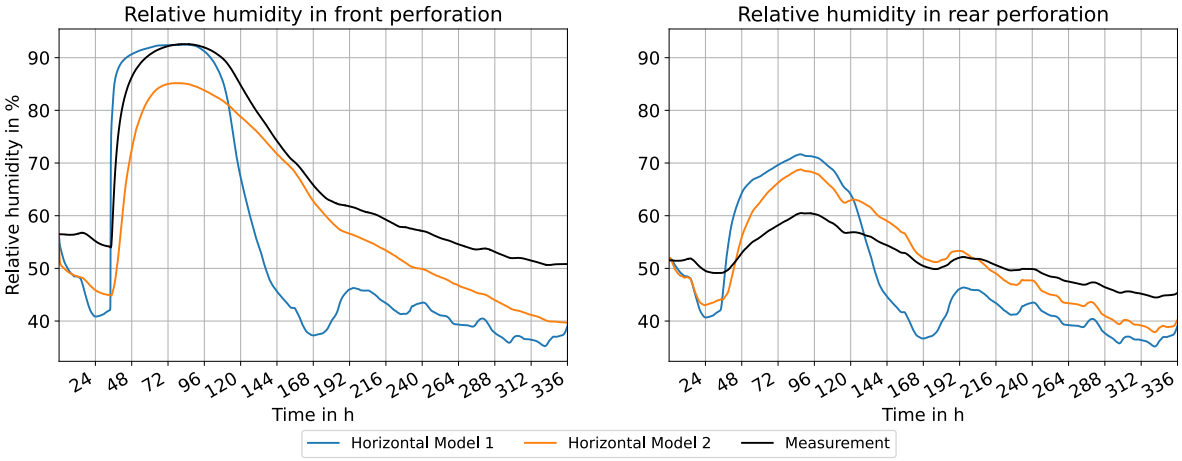


Figure 4: Relative humidity in front and rear perforation of Horizontal Models

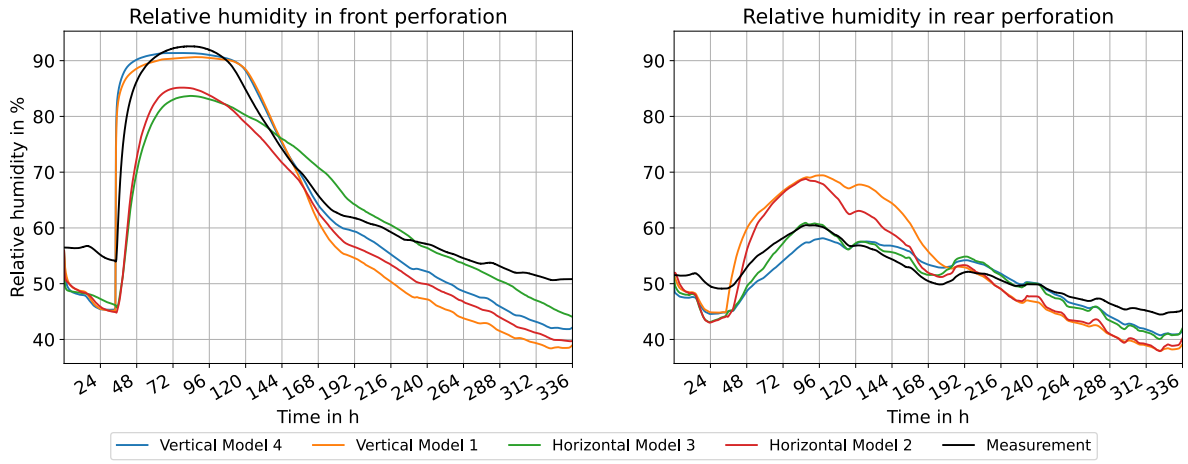


Figure 5: Relative humidity in front and rear perforation of Models with and without extra mortar

highlights clear differences in moisture behavior. Model 1 shows a significantly faster decrease in relative humidity and aligns more quickly with the ambient conditions. It also absorbs moisture more rapidly in the front perforation compared to Model 2, which exhibits a more delayed uptake and slower redistribution. These differences are also quantitatively reflected in the RMSE values over the entire duration of the simulation: in the front perforation, Model 1 reaches a RMSE of 16.09 %, while Model 2 performs substantially better with a value of 8.21 %. This trend continues in the rear perforation, where Model 1 has an RMSE of 8.85 %, and Model 2 again shows a lower deviation of 4.96%. The slower convergence of Model 2 to the ambient conditions indicates a more realistic moisture dynamics profile. These results underline the significant impact of the mortar on the internal moisture distribution within the clinker brick. Model 2, with the mortar strip, consistently performs better in replicating experimental moisture trends and is therefore selected as the reference model for subsequent investigations.

4.3 Excess mortar filling of perforations

Horizontal model 2 (section just below the pipe) already demonstrated that mortar has a major influence on moisture transport and must not be neglected. Also the significant penetration of mortar into the clinker (section 3) was not previously included in the models. Based on this observation, variants of the best models were developed that incorporate additional mortar (Vertical Model 4 and Horizontal Model 3).

The results of these models show improved agreement with the measured data, especially for the rear perforation. This improvement is illustrated in Figure 5 and also reflected in the RMSE values shown in Table 1. The only exception is the absorption phase in the front perforation where the RMSE got slightly higher due to dampened maximum values.

It can also be seen that the moisture in the front perforation rises too quickly in the vertical model and more slowly in the horizontal model. In the experiment, the actual increase in moisture lies between these two simulation results. This suggests that neither model can fully capture the dynamic moisture ingress in this zone, emphasising the need for further refinement.

Overall, these findings underscore the importance of accurately representing the actual mortar distribution within masonry for reliable hygrothermal simulation results. It can also be observed that

Table 1: RMSE-values of Models with and without extra mortar

Model	RMSE Absorption Phase Front Perforation	RMSE Drying Phase Front Perforation	RMSE Absorption Phase Rear Perforation	RMSE Drying Phase Rear Perforation
Vertical Model 1	6.33 %	8.37 %	7.33 %	6.47 %
Vertical Model 4	7.08 %	4.91 %	4.20 %	2.20 %
Horizontal Model 2	12.06 %	13.67 %	6.46 %	4.59 %
Horizontal Model 3	13.67 %	4.17 %	2.78 %	2.09 %

the humidity value slightly increased at the end of the simulation due to the addition of excess mortar. This indicates that adding excess mortar to the model could be a possible avenue to address a too rapid convergence towards ambient values.

5 Conclusion

The conducted simulations, in combination with experimental measurements, offer valuable insights into the hygrothermal behavior of perforated clinker bricks under controlled boundary conditions. A key finding is that the vertical model demonstrates a superior capacity for moisture uptake. In contrast, the horizontal model shows better performance in the redistribution of moisture. These findings underline that the choice between vertical and horizontal sectioning is not trivial, as it directly affects the predicted moisture distribution and drying behavior.

Another key result is the significant influence of mortar—both between the bricks and, critically, within the perforations. The integration of mortar residues in the perforations, notably improved simulation accuracy in the rear perforation. This highlights the importance of accounting for realistic construction effects in hygrothermal modeling. Simplified mortar-free perforations tend to underestimate or misrepresent moisture transport pathways.

While the study provides strong indications regarding model orientation and material effects, it is based on a single experimental dataset. As a result, the generalizability of the findings is limited. Moreover, certain modeling artifacts—such as the unrealistically rapid adjustment of simulated moisture and temperature levels to ambient conditions—could not yet be resolved. This highlights the need for further refinement of material parameterization and boundary condition modeling.

In future work, a 3D model should be developed to allow for a more comprehensive comparison of moisture transport and distribution within the materials. Furthermore, the simulation results should be validated using a demonstration object exposed to real weather conditions, in order to assess the practical applicability and robustness of the model under realistic environmental influences.

Acknowledgements

Parts of this work are derived from the master thesis of Andrea Thiele. The authors thank the Fraunhofer Institute for Building Physics (IBP) for providing her with a student's license of the WUFI 2D software.

Other parts of this work are derived from the "ReFaTEk" research project (promotional reference number 03EN6018A). The authors gratefully acknowledge the financial support of the German Federal Ministry for Economic Affairs and Energy within this project.

The authors would like to thank the Institute of Building Materials Research (ibac) at RWTH Aachen University for their support, particularly for conducting the analysis of the moisture storage function used in this study.

References

[1] F. U. Vogdt and M. Helbach, “Eine thermisch aktiv gedämmte, monolithische Außenwand aus Infraleichtbeton als nachhaltiges, zukunftsfähiges Wandsystem?”, *Bauphysik*, vol. 37, DOI: 10.1002/bapi.201510005

[2] M. Ibrahim, E. Wurtz, P. H. Biwole, P. Achard, and H. Sallee, “Hygrothermal performance of exterior walls covered with aerogel-based insulating rendering”, *Energy and Buildings*, vol. 84, DOI: 10.1016/j.enbuild.2014.07.039

[3] https://refatek.de/.

[4] H. M. Künzel, “Verfahren zur ein- und zweidimensionalen Berechnung des gekoppelten Wärme- und Feuchtetransports in Bauteilen mit einfachen Kennwerten”, Ph.D. dissertation, Universität Stuttgart, Stuttgart.

Overcoming Stakeholder Knowledge Gaps: A HITL-Oriented Ontology Mapping Framework for the Construction Industry

Jakob Martin[1], Timur Weilbach-Eyüboglu[1] and Vedant Girish Dalvi[1]
[1]University of Applied Sciences, Munich, Germany
E-mail(s): jakob.martin0@hm.edu, timur.weilbach-eyueboglu@hm.edu, vedant_girish.dalvi@hm.edu

Abstract: Semantic interoperability in construction and Facility Management (FM) processes is crucial for efficient information exchange, yet automated semantic matching systems often struggle with context-specific terminology and dynamic project requirements. This paper presents an extended Human-in-the-Loop (HITL) component integrated into an existing semantic matching platform for construction and FM workflows. Unlike traditional approaches with simple feedback loops, our system evolves the HITL paradigm into a role-based, domain-specific expert system. The HITL extension introduces multiple innovations: role-weighted feedback, where inputs from users (e.g., architect, engineer, facility manager) are weighted by expertise and project role; adaptive interface modes tailored to user expertise and project phase; an active learning module that detects uncertainty in match suggestions and queries experts; a conflict escalation hierarchy to resolve contradictory feedback through tiered expert intervention; and phase-specific weighting logic adjusting the matching algorithm to the current project phase. We detail the design and implementation of these features and report an evaluation of real-world construction data. Results show that the enhanced HITL system improves matching accuracy (precision 91.8%, recall 92.0%) and reduces manual corrections by 30%, demonstrating the effectiveness of integrating human expertise into semantic matching for the Architecture, Engineering, Construction and Facility Management (AEC/FM) domain.

Keywords: Human-in-the-Loop, Semantic Data Mapping, Role-Aware Interaction, Active Learning, Building Information Modeling (BIM)

DOI: 10.18154/RWTH-CONV-254909. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Semantic data integration is a longstanding challenge in the Architecture, Engineering, Construction and Facility Management (AEC/FM) domain [1]. Projects involve heterogeneous data, leading to mismatches and silos. [2]. Semantic matching systems address this by automatically identifying equivalent or related concepts across different representations, for instance, matching a facility management (FM) task description to a Building Information Modelling (BIM) component. While such

systems leverage ontologies and machine learning (ML) to improve interoperability, they often fail to capture project-specific nuances and evolving domain semantics.

Human-in-the-Loop (HITL) strategies have emerged as a promising means to address these shortcomings by integrating human expertise into automated processes [3]. However, most existing HITL implementations in the AEC/FM sector are limited to simple feedback confirmation loops, offering little contextual differentiation [4]. They do not account for the varying expertise of stakeholders, project lifecycle stages, or feedback conflicts.

This paper presents an extended HITL framework that addresses these limitations by incorporating role-specific feedback weighting, adaptive user interaction, and active learning strategies into a semantic matching system for construction and FM workflows. The framework aims to align automated suggestions more closely with project realities while reducing manual correction effort.

The remainder of this paper is structured as follows: Section 2 reviews related work in semantic matching and HITL systems. Section 3 outlines the design of the extended HITL framework. Section 4 describes the implementation within a semantic BIM matching tool. Section 5 presents evaluation results and discusses key findings. Section 6 concludes the paper and suggests future work.

2 Related Work

Prior research highlights how semantic enrichment improves information exchange and analysis across heterogeneous construction data sources [2]. Despite these advances, systems remain mostly automated and often miss domain-specific nuances. Generic algorithms struggle with varied terminology and subtle distinctions, such as between ”commissioning” and ”handover.” Some frameworks support manual validation, but usually as one-time interventions rather than continuous learning loops [2], [5].

By contrast, the broader machine learning field has developed more sophisticated HITL strategies. Active learning techniques help the system identify cases with high uncertainty and request input from human users for clarification. In contrast, interactive machine learning focuses on refining the model step by step through ongoing user feedback [3]. These techniques are especially effective in domain-specific applications like knowledge graph construction [6]. Some AEC/FM tools support compliance checks or ontology editing but still rely on simple approvals and ignore structured feedback. Expert systems in construction typically encode static rules; however, modern HITL approaches offer a more dynamic alternative, incorporating expert feedback into ongoing model updates. However, role-weighted feedback remains underexplored. While access control is common, few systems differentiate the influence of input based on expertise.

Another key dimension is user interface design. Human-centered computing has shown that adaptive User-Interfaces (UIs) help balance usability for both experts and non-experts [7]. In BIM and FM systems, overloaded interfaces are a known challenge [7]. Some tools offer role-specific views, but they are rarely dynamic. Active learning and uncertainty estimation are well-established in text classification, but still emerging in construction informatics. By learning from individual feedback and modeling how consensus is formed, such systems can capture distributed domain knowledge more effectively.

3 Methodology: Extended HITL System Design

This section outlines the methodological foundation of the extended HITL framework.

The framework overcomes limitations of conventional feedback in semantic matching. Traditional approaches typically treat all user feedback as equal and static, without considering contextual factors such as user expertise, project phase, or the confidence level of a match. In contrast, our system introduces multiple mechanisms to ensure that human input is contextually weighted, intelligently solicited, and effectively applied.

Because roles vary across procurement and delivery strategies, the framework avoids a fixed hierarchy of user influence. It uses a configurable role-weighting matrix with default profiles that can be adapted to project-specific structures.

A central component of the system is the role-based feedback weighting logic. User input is evaluated in light of the contributor's domain expertise and responsibilities. For example, during the design phase, feedback from architects or domain specialists is prioritized, while in the operational phase, facility managers receive greater influence. This weighting mechanism reflects real-world trust hierarchies and aligns the system with project workflows. Probabilistic models such as Bayesian updating and learning-to-rank methods adjust future match confidence based on the weighted feedback.

Building on this, the system adapts its interaction logic. Interface modes are selected according to user profile, technical background, and behavioral patterns such as feedback frequency or correction history. Following progressive disclosure principles [8], expert users are offered detailed editing capabilities, while non-experts receive guided, step-by-step interaction modes. If a less experienced user frequently encounters low-confidence suggestions, the system automatically shifts to Guided Mode. Conversely, repetitive confirmations may activate a Passive Mode to reduce cognitive load.

To focus expert input where it matters most, the system includes an active learning module. Matches associated with low confidence or semantic ambiguity are flagged and routed to relevant users based on role and expected information gain. This ensures that human attention is directed to critical decisions and allows the model to iteratively improve in high-impact areas.

In collaborative scenarios where user feedback conflicts, a conflict resolution mechanism is activated. The system categorizes disagreements based on severity and triggers escalation workflows if necessary. Initial strategies include prompting for clarification or requesting justification. If conflicts persist, feedback is escalated to a senior expert or resolved through a structured consensus process. All feedback and decisions are logged to support transparency and continuous learning.

Finally, the system accounts for the temporal dynamics of construction and FM projects. Semantic interpretation and feedback relevance change throughout the project lifecycle. For instance, broader semantic matches may be acceptable in early design, while stricter definitions are necessary during commissioning. Accordingly, the HITL framework adapts its weighting logic, interface behavior, and match acceptance thresholds in alignment with the current project phase.

By combining these features, the extended HITL framework strengthens a Bidirectional Encoder Representations from Transformers (BERT)-based semantic matching core with expert-guided refinement.

Rather than treating feedback as uniform, the system dynamically adapts to user role, experience, and project context. This approach ensures precise calibration of similarity scores, reduces manual correction effort, and increases trust in automated suggestions.

The overall system aims to improve interoperability by leveraging structured human input in uncertain, conflict, or phase-dependent ambiguity situations. Figure 1 illustrates the feedback and decision flow of the HITL framework, showing when human intervention is triggered and how validated input feeds back into the matcher.

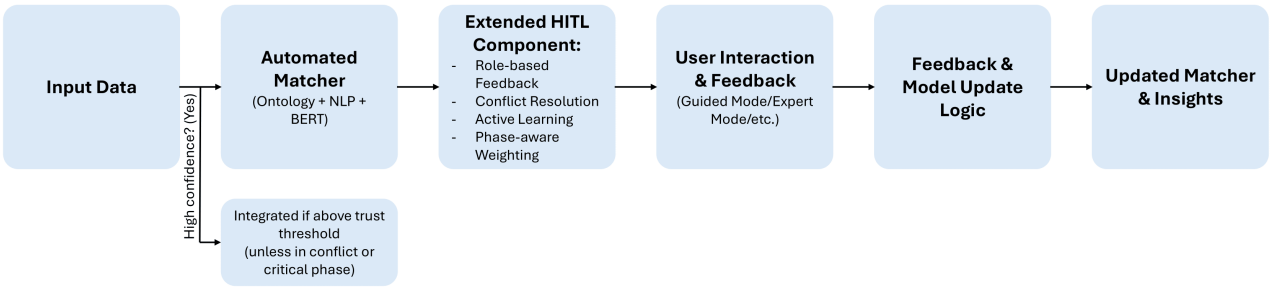


Figure 1: Processing flow of the extended HITL system. Matches with high confidence bypass human review unless flagged as critical. Ambiguous cases are routed through the HITL logic for role-aware feedback and adaptive interaction.

3.1 System Overview and Feedback Architecture

The extended HITL system builds on an existing semantic matcher that processes inputs such as process descriptions and BIM element metadata. Using ontological reasoning and natural language processing techniques, the matcher generates candidate alignments with associated confidence scores, which are then passed to the HITL module for further evaluation. These matches and associated confidence scores are passed to the HITL module for further evaluation.

The HITL module consists of an adaptive interface, a role-aware feedback weighting mechanism, and a conflict resolution system. High-confidence matches are accepted automatically, while ambiguous cases trigger human review. Feedback is evaluated based on user roles stored in a profile registry, which informs interface behavior and input weighting.

3.2 Learning Strategy and Lifecycle Adaptation

The HITL system incorporates targeted learning mechanisms and structured conflict resolution to improve semantic alignment in areas where automation proves insufficient.

The system identifies uncertain or ambiguous matches using calibrated confidence scores. Low-confidence matches are flagged and routed to qualified users. Feedback updates the model incrementally by refining similarity scores or ontology links. Prompts are embedded into task-specific interfaces at key decision points. The system is tuned to avoid overload by prioritizing high-impact interactions and suppressing redundant or low-confidence suggestions.

When users submit contradictory feedback or challenge a high-confidence suggestion, the system categorizes the disagreement by severity. Minor conflicts are resolved through clarification, while major conflicts involving users with high authority trigger an escalation procedure. In escalated cases,

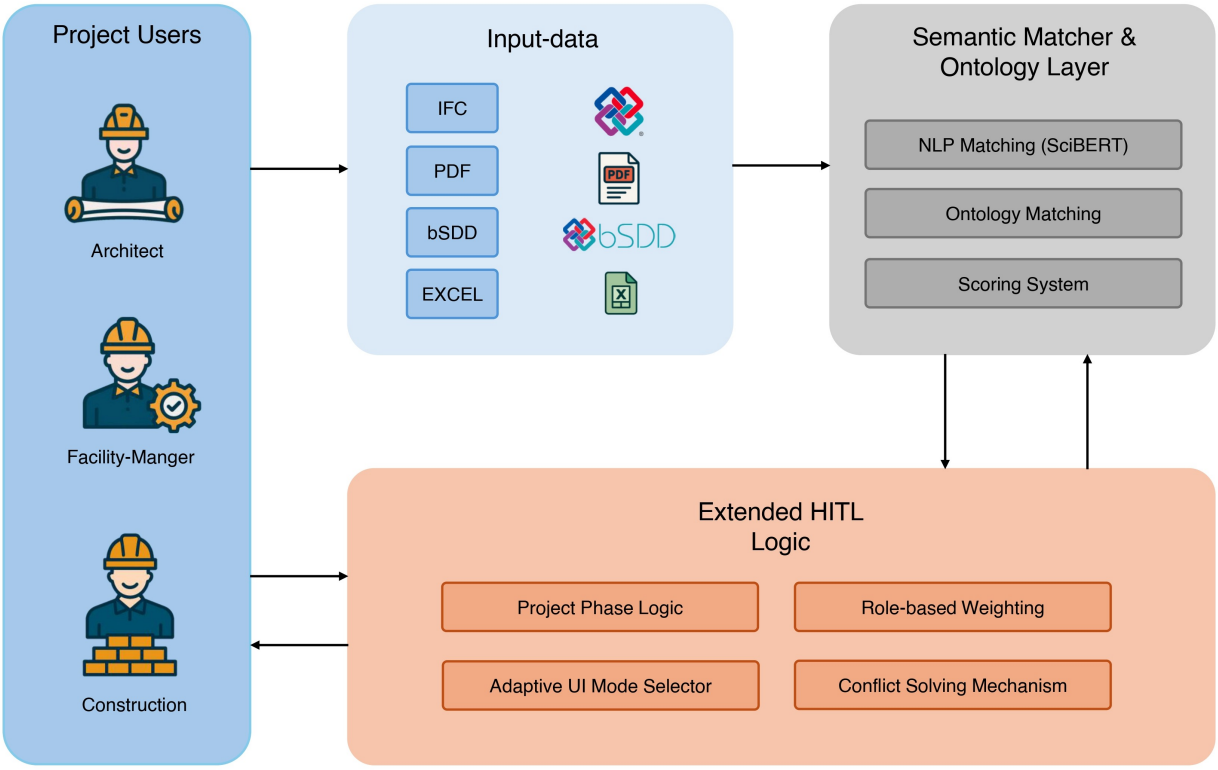


Figure 2: System architecture showing the interaction between project users, input data formats, the semantic matcher, and the HITL logic. The HITL module integrates feedback strategies tailored to user role, interface mode, conflict status, and project phase.

the system initiates second-level review, requests justification for feedback, or activates a consensus meeting mode to support collaborative resolution when necessary. All decisions are logged and used to inform future model behavior. These strategies are supported by a conflict resolution layer grounded in project hierarchies and informed by concepts from crowd annotation theory and escalation protocols, enabling structured and replicable disagreement handling.

Construction and FM projects evolve through distinct lifecycle phases with different data characteristics and stakeholder priorities. The system adjusts both algorithmic thresholds and role influence accordingly. In early design, broader matching is tolerated, and input from planners and designers is prioritized. During construction, execution feedback from roles like site managers gains importance in identifying practical constraints. In operations, stable datasets and performance monitoring elevate the role of facility management. Weight profiles are manually defined or inferred from metadata and blended during phase transitions. Similarly, the user interface and prompting strategy shift from exploratory guidance in the early stages to validation-focused input in the later stages.

4 Implementation

The extended HITL framework was integrated into an existing semantic matching platform for BIM and FM workflows [4]. The original system provided NLP-based similarity and ontology support; we added HITL components without changing the core.

The semantic matcher generates similarity scores using a combination of BERT embeddings, domain-specific patterns, and syntactic cues. These scores are processed by a scheduler that identifies low-confidence matches and generates review tasks. Tasks are routed to users by their role profiles stored in a central registry.

User feedback is stored together with the associated role information, enabling the system to prioritize corrections according to contextual authority. Submitted input incrementally updates similarity scores, category preferences, and ontology relations. All feedback actions are logged and visible to moderators. In the case of contradictory feedback, the affected match is marked as “in conflict” and displayed on a dedicated dashboard. Reviewers can then examine the feedback history, initiate clarification workflows, or issue binding decisions. If necessary, escalation rules notify higher-level roles for resolution.

The system also tracks the current project phase, which influences confidence thresholds and feedback interpretation. When phases change, new configurations are applied gradually to avoid abrupt changes in system behaviour.

To operationalize these mechanisms, the HITL module follows a recurring three-step decision cycle. First, model outputs are analysed and matches with low confidence or high uncertainty are flagged. Second, these cases are routed to users according to predefined role-based weighting factors, ensuring that the most relevant expertise is applied. Third, validated feedback is incorporated into the system through incremental weight updates, which in turn refine future suggestions. This cycle is implemented using modular scheduling components, making the process transparent and reproducible.

An event-driven architecture distributes feedback, updates, and conflicts, ensuring modularity. Prompts are context-aware and adapt to user roles: expert users can make detailed adjustments, whereas less experienced users receive guided suggestions. Tasks are prioritised by expected information gain and presented in small batches to avoid cognitive overload. All matching and feedback processes run asynchronously, allowing the system to continue processing while waiting for expert input.

5 Discussion

We focus on precision, recall, and error reduction; detailed metrics are in [4].

To assess the impact of the extended HITL framework, we compared it against a previously evaluated baseline system, which used a combination of pattern matching, lexical similarity, and a basic feedback mechanism. This baseline system was tested in a facility management scenario using 275 technical attributes across four domains (thermal, chemical, mechanical, geometric) and achieved an initial precision of 86.6 percent [4].

Table 1: Comparison of Baseline and HITL-Enhanced System Performance

Metric	Baseline System	With HITL Extension
Overall Precision	86.6%	91.8%
Overall Recall	81.4%	92.0%
Incorrect Matches	37 (13.4%)	26 (9.5%)
Avg. Confidence Score	0.79	0.87

Precision refers to the proportion of suggested matches that are correct, while recall measures the proportion of correct matches that the system successfully retrieves.

Building on the same dataset and architecture, the enhanced version incorporated role-weighted feedback, active learning, and phase-specific adaptation. As a result, the system achieved a recall of 92 percent and reduced incorrect matches in the final output by approximately 30 percent. These improvements were mainly driven by targeted human corrections and the resolution of conflicting feedback.

The active learning component played a central role by focusing user attention on uncertain cases. As the model integrated validated input over time, the number of required feedback queries decreased steadily, reducing user workload and improving efficiency. In contrast, the baseline HITL setup lacked prioritization and repeatedly prompted users for similar cases, resulting in a greater need for manual intervention.

Role-based feedback weighting also proved effective for resolving disagreements. In most cases, decisions aligned with high-authority input and matched expert ground truth. Escalation ensured corrections where needed.

Despite these positive results, several limitations and practical challenges remain.

Role-based weighting depends on accurate role attribution, which may be limited in heterogeneous or loosely structured environments.

Active learning and uncertainty estimation are well-established in text classification, but still emerging in construction informatics.

Overall, the evaluation demonstrates the potential of the proposed HITL framework to enhance semantic matching in complex construction workflows. At the same time, it highlights the need for continued refinement of role management, user interfaces, and feedback mechanisms to ensure robustness and usability across diverse project environments.

6 Conclusion and Future Work

This study shows that HITL enhances semantic matching in the AEC/FM Industry. By incorporating role-sensitive feedback, active learning, and phase-aware adaptation, the framework enables targeted expert input and improves both precision and robustness over purely automated methods.

Key contributions include a modular feedback architecture, context-aware interaction modes, and a conflict resolution mechanism grounded in domain authority. These components embed expertise without overwhelming users.

Future work will focus on several open challenges identified during evaluation. One priority is to improve the system's resilience in projects with inconsistent or missing role metadata, for example by exploring automatic role inference from behavior patterns or interaction history. Another direction is to enhance uncertainty estimation in sparse data domains through hybrid confidence metrics that combine semantic distance, data quality, and user correction history. Additionally, more adaptive user interfaces could improve accessibility, particularly for users with limited technical expertise.

7 Data availability statement

The data supporting this study's findings are available at https://gitlab.lrz.de/jmartin/HITL_extension_nlp_matching_tool, ensuring that they are Findable, Accessible, Interoperable, and Reusable (FAIR).

Acknowledgements

This research was supported by the Nemetschek Innovation Foundation and benefited from collaboration with industry partners within the framework of the openDBL project. The openDBL project is funded by the European Union's Horizon Europe research and innovation program under grant agreement No. 101092161, providing valuable test cases during development.

References

[1] P. Pauwels, "Supporting decision-making in the building life-cycle using linked building data", *Buildings*, vol. 4, no. 3, pp. 549–579, 2014. DOI: 10.3390/buildings4030549

[2] T. Bloch, "Connecting research on semantic enrichment of BIM – review of approaches, methods and possible applications", *Journal of Information Technology in Construction (ITcon)*, vol. 27, pp. 416–440, Apr. 2022. DOI: 10.36680/j.itcon.2022.020

[3] D. R. Honeycutt, M. Nourani, and E. D. Ragan, "Soliciting Human-in-the-Loop User Feedback for Interactive Machine Learning Reduces User Trust and Impressions of Model Accuracy", in *Proceedings of the Eighth AAAI Conference on Human Computation and Crowdsourcing (HCOMP-20)*, Virtual Conference: AAAI Press, 2020. [Online]. Available: https://ojs.aaai.org/index.php/HCOMP/article/view/7464

[4] J. Martin, C. Preidel, S. Vilgertshofer, and A. Borrmann, "Enhancing Construction Data Integration through Dynamic Ontology Alignment and Automated Attribute Mapping", in *Proceedings of the European Conference on Computing in Construction (EC³)*, Porto, Portugal: EC3 / CIB W78, Jul. 2025. [Online]. Available: www.dx.doi.org/%2010.35490/EC3.2025.250

[5] Z. Dragisic, D. O. Faria, C. Pesquita, J. Euzenat, and V. Svátek, "User validation in ontology alignment", *Journal of Web Semantics*, vol. 43, pp. 90–106, 2017. DOI: 10.1016/j.websem.2016.11.001

[6] M. Schröder, C. Jilek, and A. Dengel, "A Human-in-the-Loop Approach for Personal Knowledge Graph Construction from File Names", in *Proceedings of the Third International Workshop on Knowledge Graph Construction at the 19th Extended Semantic Web Conference (KGCW–ESWC 2022)*, ser. CEUR Workshop Proceedings, vol. 3141, Hersonissos, Greece: CEUR-WS.org, 2022. [Online]. Available: https://ceur-ws.org/Vol-3141/paper2.pdf

[7] J. Hussain et al., "Model-based adaptive user interface based on context and user experience evaluation", *Journal on Multimodal User Interfaces*, vol. 12, no. 1, pp. 1–16, Feb. 2018. DOI: 10.1007/s12193-018-0258-2 [Online]. Available: https://doi.org/10.1007/s12193-018-0258-2

[8] E. Mosqueira-Rey, E. Hernández-Pereira, D. Alonso-Ríos, J. Bobes-Bascarán, and Á. Fernández-Leal, "Human-in-the-loop machine learning: a state of the art", *Artificial Intelligence Review*, vol. 56, no. 4, pp. 3005–3054, Aug. 2022. DOI: 10.1007/s10462-022-10246-w

Targeted Data Exchange for Efficient MEP Openings in BIM using IDS

Timur Weilbach-Eyüboglu[1], Jakob Martin[1] and Vedant Girish Dalvi[1]
[1]Hochschule München University of Applied Sciences, Munich, Germany
E-mail(s): timur.weilbach-eyueboglu@hm.edu, jakob.martin0@hm.edu, vedant_girish.dalvi@hm.edu

Abstract: Effective coordination of openings for Mechanical, Electrical, and Plumbing (MEP) systems within structural models remains a significant hurdle in modern Building Information Modeling (BIM) processes, frequently resulting in project inefficiencies and rework. Current practices rely on provisional placeholders lacking direct semantic and geometric linkage to host elements, necessitating tedious manual transformations. This paper introduces a demand-driven framework that automates the conversion of provisional voids into semantically structured IfcOpeningElement instances – complete with system associations, preserved design attributes, and precise spatial integration within host elements – thereby enhancing cross-disciplinary coordination and model fidelity in BIM workflows. Combining Information Delivery Specification (IDS) with advanced spatial reasoning capabilities facilitates precise extraction and mapping of non-geometric attributes and spatial characteristics. The methodology enhances the automated reconciliation of design intent across disciplines. This work contributes a concrete method for automating cross-disciplinary coordination of openings in BIM, improving the reliability of MEP-structure integration and reducing manual modeling overhead in data-rich, federated models.

Keywords: Building Information Modeling (BIM), MEP Coordination, Information Delivery Specification (IDS), Automated Workflows

DOI: 10.18154/RWTH-CONV-254910. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Coordinating Mechanical, Electrical, and Plumbing (MEP) systems with structural elements represents a complex challenge in contemporary construction projects [1]. The precise planning and execution of openings and breakthroughs – where MEP services such as Heating, Ventilation, and Air Conditioning (HVAC) ducts, electrical conduits, and plumbing pipes must penetrate structural components like walls, slabs, and beams – has been a significant source of coordination difficulties, costly on-site modifications, and substantial project delays. Traditional coordination approaches, which rely heavily on 2D drawings and manual work across disciplines, have proven inadequate for managing the increasing complexity of modern building systems. [2], [3]

The advent of Building Information Modeling (BIM) has fundamentally transformed this reactive problem-solving paradigm into a proactive coordination approach [1], [4]. Model-based clash detection

enables the automated identification of critical geometric overlaps between discipline-specific models during the planning phase, reducing the need for time-consuming coordination efforts in later project stages.

However, despite these advances, the coordination of MEP openings remains hindered by a fundamental modeling constraint: MEP engineers cannot directly create *IfcOpeningElements* – the designated BIM elements for physical openings – in their authoring environments because *IfcOpeningElements* require a host component, such as walls, slabs, or beams. These structural elements typically originate from architectural or structural models, so provisional placeholders for openings are added to the structural Industry Foundation Classes (IFC) models by architects or structural engineers. As a result, MEP engineers cannot create *IfcOpeningElements* themselves and must rely on these placeholders to represent openings in their models. To coordinate openings, MEP planners use *IfcBuildingElementProxy* elements with the *PredefinedType ProvisionForVoid* (*BEP-PV*) as volumetric placeholders to express their breakthrough requirements. These proxies communicate the intended size, shape, and position of openings and carry MEP-specific metadata such as system assignments, identification numbers, and dimensions. However, this approach introduces a critical limitation: they lack the semantic links to host elements required for downstream processing and implementation. This deficiency is a well-recognized problem within the industry. *BuildingSMART* has highlighted the need for automated workflows that convert *BEP-PV* placeholders into proper *IfcOpeningElements* while preserving critical attributes and ensuring correct spatial integration within host elements [5]. Existing solutions, as discussed in the following *State of the Art* section, rely on manual post-processing, which are misaligned with industry demands for efficient, selective, and software-neutral data exchange processes.

This paper addresses these limitations by proposing a framework for MEP opening coordination. Expanding our prior work [6], the proposed framework enables the automated transformation of *BEP-PV* proposals into properly coordinated *IfcOpeningElements*. Although the presented approach does not automate this process at the authoring software level, it offers a structured and reproducible framework that integrates the resulting openings within structural models with minimal manual effort. This contribution supports *buildingSMART's* automation goals and promotes more scalable, open, and interoperable BIM workflows across software environments.

2 State of the Art: IFC-Based Coordination of Openings

In current BIM workflows, openings and penetrations are coordinated by sharing domain models via IFC exports. Typically, MEP engineers insert *Provision for Void* placeholders in the architectural or structural model to reserve space for openings. In IFC, this is implemented using a *BEP-PV*. The IFC schema describes this usage explicitly: a building element proxy "as a provision for voids" represents a requested volume of space, which may later become an actual opening [7].

Such proxies carry a solid swept volume (matching or exceeding the thickness of the host element) and can be annotated with properties (via *Pset_BuildingElementProxyProvisionForVoid*) for shape, size, or system information. For example, an MEP duct might be represented by a cylindrical proxy with properties for diameter and required depth linked to a system ID in the property set. When

imported into other platforms (e.g., structural BIM or clash-detection tools), the *BEP-PV* appears as a generic proxy object that flags a requested opening.

Ultimately, the architect or structural engineer must convert these placeholder volumes into real openings in the building elements. In IFC, this is done using *IfcOpeningElement* (or *IfcOpeningStandardCase* for prismatic holes) inserted into the host via an *IfcRelVoidsElement* link. An *IfcOpeningElement* "represents a void within any element that has a physical manifestation". It is assigned a *PredefinedType* of *OPENING* or *RECESS* and may carry its own property set (e.g., *Pset_OpeningElementCommon* for purpose, status, fire rating, etc.) [8].

In practice, however, this conversion is not automated: designers must create these openings manually. ArchiCAD, for example, imports *BEP-PV* proxies as generic Morph objects flagged as provisions for void; users must then manually convert each Morph object into an ArchiCAD opening. Only after this manual step does the element become an *IfcOpeningElement* tied to its host upon export. Revit requires a similar manual acceptance of the void request [9].

The handling of *BEP-PV* across common tools remains fragmented. ArchiCAD provides a "Convert Provision for Void" command but lacks full automation. Revit's IFC export includes *BEP-PV* elements, yet offers no native workflow for automatically updating openings. Tekla Structures can export voids as *IfcOpeningElement* or generic subtractions but does not interpret imported *BEP-PV* to generate openings. Coordination platforms like Navisworks or Solibri visualize *BEP-PV*, but do not translate them into actionable geometry.

Due to these limitations, current workflows remain predominantly supply driven. Teams exchange full IFC models to communicate opening requests, which introduces inefficiencies and increases the risk of errors. Semantic information, such as system assignments or dimensional properties, is often lost during manual reimplementation. BuildingSMART Germany explicitly calls for the "automated conversion of opening proposals (ProvisionForVoid) into real openings including automatic labeling and transfer of properties" [5].

In summary, IFC-based opening coordination today relies on *BEP-PV* proxies as placeholders and manual conversion to *IfcOpeningElement*s. Automation is minimal, and interoperability suffers due to redundant modeling efforts and semantic loss. There is a pressing need for improved tooling and structured, demand-driven workflows to bridge the gap between placeholder elements and semantically rich, host-integrated openings.

3 Methodology

Building on our earlier work [6], which introduced a *pull-based* workflow concept using Information Delivery Specification (IDS), this paper further develops and operationalizes the approach within a targeted application context for MEP opening coordination. Our methodology addresses the established coordination gap by extending the traditional IDS-driven *pull* approach with geometric reasoning capabilities, enabling the automated transformation of *BEP-PV* into semantically and geometrically compliant *IfcOpeningElements*.

The *pull-based* method enables information requesters to actively articulate data needs using IDS, which serve as transport containers and validation mechanisms for targeted data extraction, as

elaborated in [6]. This demand-driven process ensures that only relevant information is retrieved and integrated, reducing coordination cycles and improving data quality across discipline-specific models. The IDS framework operates through two primary components: the *Applicability* criteria that establish the specification's scope by identifying relevant IFC entities and element types, and the *Requirements* that articulate specific data constraints, encompassing mandatory attributes, acceptable value ranges, and formatting standards. For our opening coordination workflow, IDS specifications define the semantic requirements for *BEP-PV* elements, including their classification, dimensional properties, system assignments, and identification parameters.

However, a critical limitation of IDS is its inherent inability to represent or process geometric data. While it excels at specifying semantic requirements, it cannot capture spatial characteristics essential for opening coordination, such as precise positioning, geometric relationships, and spatial intersections with host elements. To address this limitation, our methodology integrates a geometric feature extraction pipeline alongside the IDS-based semantic validation process. This dual-processing approach enables comprehensive handling of both semantic and spatial data requirements, bridging the gap between semantic specifications and geometric implementation. The geometrical analysis process extracts spatial characteristics from *BEP-PV* elements, including bounding box coordinates, centroid positions, and volumetric properties, while the IDS framework ensures semantic compliance and attribute preservation throughout the transformation process.

Figure 1 illustrates the conceptual framework contrasting the traditional push-based workflow, where information is actively sent from the provider to recipients without a specific request, with our proposed *pull-based* approach for automated *BEP-PV* to *IfcOpeningElement* transformation.

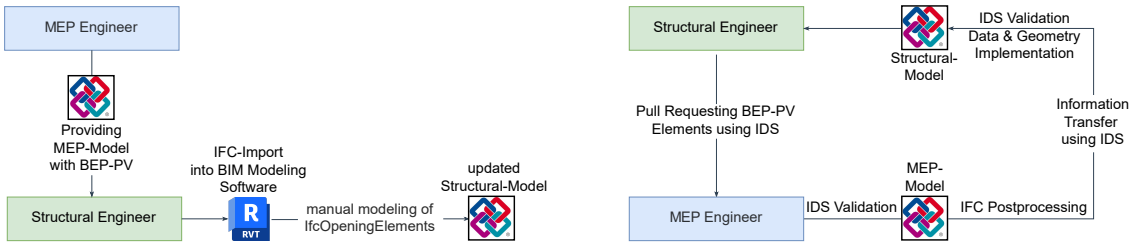


Figure 1: Comparison Push-based Approach (left) and Pull-based Approach (right)

The workflow operates through three sequential phases: semantic validation, geometric feature extraction, and spatial association. During the semantic validation phase, IDS specifications are applied to verify that *BEP-PV* elements meet predefined requirements for attributes, classifications, and property assignments. This validation ensures data quality and completeness before proceeding to geometric processing, preventing downstream errors and maintaining semantic integrity throughout the transformation process.

The geometric feature extraction phase employs computational geometry algorithms to analyze the spatial properties of validated *BEP-PV* elements. This process extracts essential geometric descriptors, including three-dimensional bounding boxes, centroid coordinates, principal axes orientations, and volumetric extents. These geometric features serve as the foundation for subsequent spatial reasoning operations, enabling precise positioning and orientation analysis within the broader model context.

The extraction process accommodates various geometric representations and coordinate systems, ensuring consistency across diverse modeling approaches and software environments.

The methodology employs a spatial reasoning approach based on bounding volume intersection analysis to associate *BEP-PV* elements with their corresponding host elements. Each provision element's extracted geometric features are systematically compared against potential host elements, including walls, slabs, beams, and other structural components. The spatial association algorithm evaluates geometric relationships through intersection testing, proximity analysis, and containment detection. When a meaningful spatial overlap is identified between a provision element and a structural component, the algorithm establishes a host relationship and prepares the data for the creation of an *IfcOpeningElement*. This automated spatial reasoning process eliminates manual host identification while maintaining accuracy through geometric validation, enabling scalable processing of large-scale opening coordination scenarios while preserving the semantic richness of the original *BEP-PV* elements throughout the transformation into compliant *IfcOpeningElements*.

4 Implementation

The already mentioned pull-based transformation of *BEP-PV* elements into *IfcOpeningElements* is implemented through a structured Python-based workflow, covering IDS creation, IFC validation, geometry extraction, collision detection, and IFC model editing. This approach ensures automation and software neutrality when coordinating MEP openings.

The process starts with generating an IDS file that filters *IfcBuildingElementProxy* entities with *PredefinedType = PROVISIONFORVOID*. The IDS includes mandatory attributes such as GUID, Depth, Height, and Width to identify and describe provision elements in the MEP IFC model. Validation is conducted using the *IfcTester* Python Package, which parses the model and extracts compliant elements based on the defined criteria [10]. These validated elements serve as the basis for downstream geometric and semantic analysis. Figure 2 illustrates the *IDS-Pull-Request* and the resulting validated *BEP-PV* element.

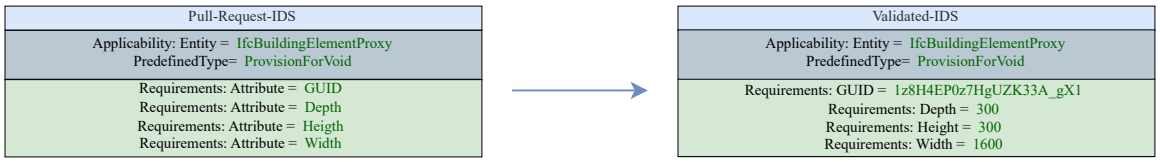


Figure 2: IDS-Pull-Request and Validated BEP-PV element

Since the IDS schema does not support spatial information, the developed software framework uses the Python Package *IfcOpenShell* to extract geometry from validated elements [11]. Each *BEP-PV* is processed to compute an axis-aligned bounding box (AABB), which simplifies spatial containment and overlap checks. The bounding box serves as a proxy for spatial reasoning, enabling lightweight yet practical geometric comparisons.

In parallel, all associated metadata, including values from *Pset_BuildingElementProxyProvisionForVoid*, is extracted to preserve the element's semantic context. The geometric and property data form a comprehensive digital representation of the void reservation, which can be compared against potential host elements.

Collision detection is performed by loading the MEP and structural IFC models simultaneously. The algorithm checks for 3D AABB intersections between provision elements and structural components like walls or slabs. Bounding boxes are evaluated along the global X, Y, and Z axes to detect spatial overlaps.

Precomputed intersections identify candidate host elements, whose corresponding GUIDs and classifications are read from the input data to associate *BEP-PV* reservations with their structural hosts. This mapping enables the placement of semantically correct and geometrically accurate openings in the model.

An exemplary *Answer-IDS*, as illustrated in Figure 3, is generated for each valid pairing to capture all relevant details: bounding box coordinates, element dimensions, source and host GUIDs, and metadata. This structured file is used for automated transformation into *IfcOpeningElements*.

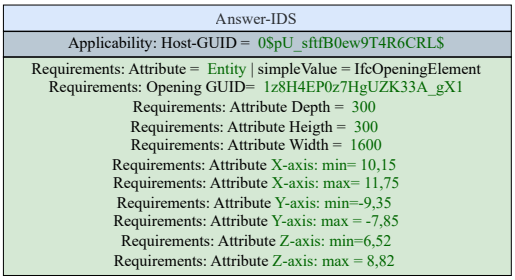


Figure 3: Information Transport - Answer-IDS

The transformation phase generates a new *IfcOpeningElement* for each validated provision, using geometry and placement derived from the Answer-IDS. Rectangular cross-sections are defined using *IfcRectangleProfileDef* and extruded with *IfcExtrudedAreaSolid* to reflect the required dimensions. Each opening is placed relative to the host's local coordinate system.

Openings are linked to host elements via *IfcRelVoidsElement* to establish the logical relationship. In cases of multi-host overlaps, appropriate relationship networks are created. These connections ensure BIM tools correctly interpret the modifications and support downstream use in coordination or quantity take-off.

The final result is a modified IFC model containing semantically rich, host-integrated *IfcOpeningElements* based on MEP provision inputs.

5 Discussion and Conclusion

This paper presents a pull-based framework for the automated transformation of *BEP-PV* elements into semantically integrated *IfcOpeningElements*, using IDS to formalize demand-driven information exchange. Combining IDS validation with spatial reasoning enables rule-based extraction and mapping of geometric and semantic data, supporting consistent and scalable coordination between MEP and structural domains.

Using IDS as a lightweight filtering and specification tool proved effective for defining information needs in a structured, machine-readable format. This improves clarity in interdisciplinary communication and helps avoid information overload by delivering only relevant model data. Geometry extraction

through bounding box analysis provided a viable method for automating host-element association and placement of openings, substantially reducing manual workload.

Nonetheless, several limitations remain. The framework assumes a well-structured source model with complete semantic data, which is not always guaranteed – especially in legacy or inconsistently modeled projects. The current geometric matching approach, based on axis-aligned bounding boxes, offers simplicity but may lack precision in detecting complex or irregular openings. Enhancing the system with tolerance-based logic or semantic pattern recognition could improve robustness.

Interoperability also poses challenges. While IDS is platform-neutral, its effectiveness depends on consistent implementation and support within authoring tools. Broader adoption will benefit from standardized validation pipelines and extended IDS definitions, potentially covering aspects such as grouped systems or sequencing requirements.

Despite these constraints, the proposed workflow aligns with current trends in model-based delivery and automation in BIM.

6 Outlook

Future improvements to the *pull-based* workflow for transforming *BEP-PV* elements into *IfcOpeningElements* should focus on enhancing both geometric precision and semantic robustness – e.g., the ability to accurately capture, interpret, and maintain meaningful non-geometric information such as element classification, properties, and relationships, even when source models have incomplete or inconsistent data. A key enhancement lies in enabling partial and context-aware geometry extraction. Instead of processing entire *BEP-PV* volumes, isolating only intersecting regions with structural elements would reduce data overhead and increase the relevance of exchanged information.

Another promising direction is advancing spatial reasoning. Extending IDS or coupling it with external logic to evaluate partial overlaps, proximity thresholds, or containment hierarchies could improve host element identification and better align with real-world coordination needs.

Addressing inconsistent or missing semantic data in source models remains critical. Artificial Intelligence-assisted classification or rule-based enrichment could infer properties or system relationships, enabling more reliable transformations even in incomplete or non-standardized models.

Future work should emphasize real-world integration, broader use cases (e.g., fire safety, clash resolution), and collaboration with industry partners to validate the approach under diverse project conditions.

In summary, refining geometry extraction, enhancing reasoning mechanisms, and enriching semantic interpretation will strengthen the robustness and applicability of pull-based BIM coordination – supporting more automated, efficient, and interdisciplinary workflows.

Data availability statement

The data supporting this study's findings are available at https://gitlab.lrz.de/tweilbach/pfv-ifcopening.git, ensuring that they are Findable, Accessible, Interoperable, and Reusable (FAIR).

Acknowledgments

The authors gratefully acknowledge the support and resources provided by the Bavarian Construction Industry Association (Bayerischer Bauindustrieverband), enabling our research and its practical application.

References

[1] Y. H. Teo et al., “Enhancing the MEP Coordination Process with BIM Technology and Management Strategies”, *Sensors*, vol. 22, no. 13, 2022. DOI: 10.3390/s22134936 [Online]. Available: https://www.mdpi.com/1424-8220/22/13/4936

[2] S. Mehrbod, S. Staub-French, N. Mahyar, and M. Tory, “Beyond the clash: investigating BIM-based building design coordination issue representation and resolution.”, *Journal of Information Technology in Construction*, vol. 24, 2019.

[3] R. Sacks and R. Barak, “Impact of three-dimensional parametric modeling of buildings on productivity in structural engineering practice”, *Automation in Construction*, vol. 17, no. 4, pp. 439–449, 2008. DOI: https://doi.org/10.1016/j.autcon.2007.08.003

[4] K. Chauhan, A. Peltokorpi, H. Abou-Ibrahim, and O. Seppänen, “Mechanical, Electrical and Plumbing Coordination Practices: Case Finnish Construction Market”, in *Annual Conference of the International Group for Lean Construction*, International Group for Lean Construction (IGLC), 2022, pp. 635–644.

[5] buildingSMART Deutschland, *Schlitz- und Durchbruchsplanung auf Basis von IFC [Slot and Opening Planning Based on IFC]*, https://ucm.buildingsmart.org/de/use-cases/3031/de, 2021.

[6] T. Weilbach-Eyüboglu, C. Preidel, and S. Vilgertshofer, “Enhancing BIM Data Information Exchange processes: A Pull-Based Framework for Targeted Data Extraction and Validation”, in *Proceedings of the European Conference on Computing in Construction (EC³)*, Porto, Portugal: EC3 / CIB W78, Jul. 2025. [Online]. Available: www.dx.doi.org/10.35490/EC3.2025.329

[7] *ISO 16739-1:2018 – Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries*, International Standard, ISO, 2018.

[8] C. Eastman, P. Teicholz, R. Sacks, and K. Liston, *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors*, 3rd. Hoboken, NJ: Wiley, 2018.

[9] Q. Lu and Y. H. Wong, “A BIM-based approach to automate the design and coordination process of mechanical, electrical, and plumbing systems”, *Transactions Hong Kong Institution of Engineers*, vol. 25, no. 4, pp. 273–280, 2018. DOI: 10.1080/1023697X.2018.1537813

[10] *IfcTester — IfcOpenShell Documentation*, https://docs.ifcopenshell.org/ifctester.html, 2025.

[11] IfcOpenShell Contributors, *IfcOpenShell documentation*, IfcOpenShell Project, 2025.

Qualitative Analyse räumlicher Bedarfe und Anforderungen infolge von Homeoffice

Anna Langenbeck[1], Marc Syndicus[1], Christoph van Treeck[1] and Jérôme Frisch[1]
[1]Lehrstuhl für Energieeffizientes Bauen (E3D), RWTH Aachen University, Aachen, Deutschland
E-mail(s): langenbeck@e3d.rwth-aachen.de

Abstract: Neue Arbeitskonzepte, darunter die Nutzung des Homeoffice, finden seit der Pandemie verstärkt Anwendung. Diese Arbeit untersucht das Gesamtkonstrukt der Homeoffice-Nutzung mit dem Ziel, Nutzungsprozesse und längerfristige Auswirkungen zu identifizieren. Neue Bedarfe und Anforderungen in Bezug auf Arbeitsräumlichkeiten werden abgeleitet. Es erfolgt eine Analyse von 35 teilstrukturierten Interviews mit Funktionstragenden aus drei Unternehmen, die im Rahmen des BMWK (BMWE) finanzierten Verbundprojektes „HOffEn - Home-Office und Energiewende“ durchgeführt wurden. Unsere Ergebnisse deuten darauf hin, dass am Standort *Büro* eine Verschiebung des Tätigkeitsschwerpunktes hin zu sozialer Interaktion und Arbeit im Team stattfindet. Einzelarbeiten und digitale Kommunikation werden teils gezielt ausgelagert. Im häuslichen Umfeld ist eine Flächenbereitstellung erforderlich. Gleichzeitig bestehen Ansätze Büroflächen zu reduzieren. Anschließend werden Chancen und Grenzen gesamtheitlicher Flächeneinsparungen diskutiert.

Keywords: Homeoffice, Bürokonzepte, Nutzungsprozesse, Qualitative Interviews, Planung

DOI: 10.18154/RWTH-CONV-254911. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Einleitung

Aus der Konjunkturumfrage des ifo Instituts [1] geht hervor, dass etwa 25 % der Beschäftigen in Teilen von zu Hause arbeiten. Auch der aktuelle Ergebnisreport der Konstanzer Homeoffice Studie [2] bestätigt, dass die Mehrheit der Befragten ein hybrides Arbeitsmodell bevorzugt. Es wird geschlussfolgert, dass Unternehmen durch Anpassung an die neue Arbeitskultur langfristig betrachtet ihre Leistungsfähigkeit sichern können. In der vorliegenden Arbeit adressieren wir die Frage inwiefern nutzungsprozessbasierte Anpassungen in Bezug auf die räumliche Gestaltung von Arbeitsplätzen zu einer gewinnbringenden Nutzung von Homeoffice beitragen können. Die DGUV [3] differenziert bei Integration mehr als gelegentlich genutzter Homeoffice-Arbeitsplätze beispielbasiert zwischen funktionaler (1,92 bis 2,46 m²) und optimaler (2,88 m²) Situation. Aus dem Sozial Bericht 2024 [4] geht hervor, dass Wohnraummangel und steigende Wohnkosten drängende Probleme darstellen. Eine Überbelastung durch Wohnkosten gemäß Definition des Statistischen Amtes der Europäischen Union betraf im Jahr 2022 11,9 % der Personen in Deutschland. Krause et al. prognostizieren eine langfristige Büroflächenverkleinerung wegen Homeoffice von 12 % im wahrscheinlichen Szenario [5], sodass durch Umnutzung zu Wohnraum in den Top-7-Städten Deutschlands theoretisch rund 60 000 Wohnun-

gen entstehen könnten [6]. Potentielle Auswirkungen neuer Arbeitskonzepte auf Flächenbedarfe im häuslichen Umfeld und am Unternehmensstandort sowie daraus resultierende Verbräuche scheinen neben Gründen der Nachhaltigkeit auch aus sozialer Perspektive relevant. Damit Einsparungen nicht zu Lasten der Funktionalität vorgenommen werden, könnten allgemein zugängliche Planungstools zukünftig auch Nutzungsprozesse berücksichtigen. Die Ansätze dieser Arbeit können als Grundlage dienen. Gemeinschaftsbüros meinen im Folgenden Büros, die mehrere Arbeitsplätze beinhalten.

2 Methodik

2.1 Analyse qualitativer Interviews

Die Arbeit basiert auf 35 teilstrukturierten Interviews mit Funktionstragenden verschiedener Unternehmen, die im Rahmen des BMWK (BMWE) geförderten Projektes „HOffEn - Home-Office und Energiewende“ von August 2022 bis Januar 2023 durch die Projektpartner EBC, E3D, HLS und IÖW geführt wurden. Im Anschluss wurden die Transkripte durch das E3D codiert und ausgewertet. Ein Votum der zuständigen Ethik-Kommission wurde vor Durchführung der Interviews eingeholt. Zielgruppe der Befragung waren Führungskräfte sowie Mitarbeitende aus den Bereichen Human Resources, Mitarbeitenden-Vertretung, Gesundheit, IT, Facility-Management und Mobilität. Neben allgemeinen Fragen wurden weitere Fragestellungen gemäß der Expertise der Teilnehmenden fach- bzw. tätigkeitsspezifisch ausgewählt und in die Leitfäden integriert. Die Gespräche wurden in Form von Audios oder Notizenprotokollen dokumentiert und im Anschluss teils softwaregestützt transkribiert. Ziel der explorativen, qualitativen Herangehensweise war die Entwicklung eines tieferen Verständnisses für das Verhalten von Nutzenden im Kontext der anteiligen Arbeit aus dem häuslichen Umfeld. Die induktiv am Material gebildeten Kategorien wurden in ein hierarchisches Kategoriensystem überführt. Codiert wurden sowohl stellvertretend genannte Inhalte als auch spezifische (persönliche) Schilderungen. Die Codierung wurde anhand von drei Interviews stichprobenhaft über die Bestimmung der Intercoderübereinstimmung geprüft. Darüber hinaus erfolgte eine Reflexion hinsichtlich weiterer in Mayring [7] genannter Gütekriterien für die qualitative Forschung. Eine Einordnung der Ergebnisse wird vorgenommen. Sowohl die Codierung als auch die Auswertung (inkl. Ermittlung der Intercoderübereinstimmung) wurden in MAXQDA Plus 2022 (Release 22.7.0) durchgeführt. Weiterführende quantitative Untersuchungen wurden durch die Projektpartner vorgenommen [8].

2.2 Aufbereitung der Ergebnisse und weiterführende Schritte

Die im Rahmen des Projekt HOffEn [8] erarbeiteten Spannungsfelder bezogen auf die beiden Standorte *Homeoffice* und *Büro* werden hier aufgegriffen und in eine gemeinsame Grafik integriert (siehe Abb. 1). Zunächst werden die vorliegenden codierten Motive für die Wahl des Standort *Homeoffice* bzw. *Büro* [8] nach primär *tätigkeitsbezogenen* und primär *personenbezogenen* Gründen neu strukturiert. Die Gruppe *Sonstiges* umfasst Kategorien, deren Zuordnung aufgrund mangelnden Kontexts nicht möglich scheint. Basierend auf den ausgeführten Tätigkeiten und weiteren Maßnahmen/Ansätzen zur Bewältigung der beschriebenen Risiken/Herausforderungen, werden Bedarfe und Anforderungen an die beiden Standorte abgeleitet und mögliche Implikationen für die Flächennutzung dargelegt. Weitere vorliegende, codierte Aussagen aus den Interviews und Ergebnisse des Abschlussberichts werden bei Bedarf assoziativ hinzugezogen, um die Argumentation zu stützen [8]. Die Anzahl der Interviews,

in denen eine Kategorie oder ein Bereich genannt wurde, wird aus Gründen der Transparenz und als Indikator für präsente Aspekte angegeben.

3 Ergebnisse

Bei den im Rahmen der Interviews getätigten Schilderungen handelt es sich um subjektive, kontextgeprägte Erfahrungsberichte. Unsere Ergebnisse sind entsprechend als mögliche Ansätze und Denkanstöße für die Planung von Arbeitsräumlichkeiten zu verstehen, die es zu hinterfragen und situationsabhängig zu bewerten gilt. Um dies zu erleichtern, diskutieren wir abschließend, ob gesamtheitlich betrachtet Flächeneinsparungen möglich und sinnvoll sein können.

3.1 Spannungsfeld bei anteiliger Arbeit von zu Hause

7 der 35 Befragten (n=7) gaben an, dass aus Mitarbeitendensicht u. a. Freiwilligkeit Voraussetzung für die Arbeit von zu Hause ist. In Teams, die hybrid (*Homeoffice* und *Büro*) tätig sind, kann es zu Veränderungen der Arbeitsweise (z. B. digitale Zusammenarbeit) kommen. Standortbezogen oder die Interaktion betreffend können neue Herausforderungen entstehen, etwa das Miteinander im Team, was sich z. B. durch zunehmende Anonymität / Entfremdung zeigen kann.

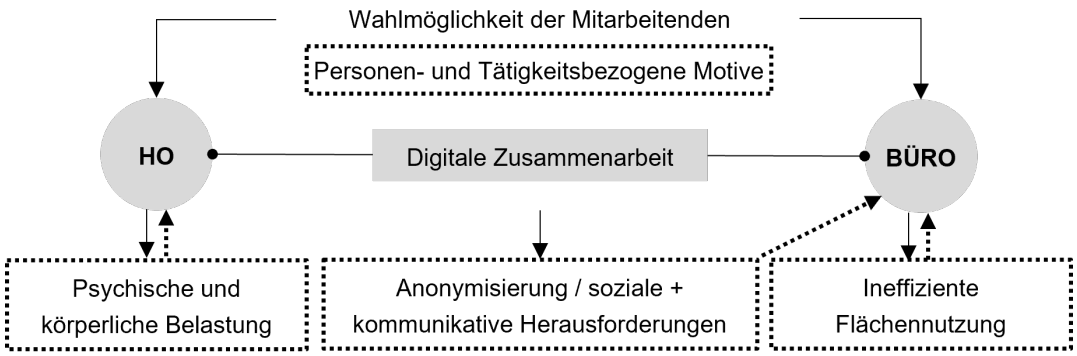


Abbildung 1: Spannungsfeld bei hybrider Arbeit von zu Hause und vom Unternehmensstandort aus

3.2 Beweggründe bei der Wahl des Arbeitsortes und dessen Nutzung

Bei der Untersuchung der Motive für die Wahl des Standort *Homeoffice* bzw. *Büro* wurden die im Rahmen des Projektes spezifisch für den jeweiligen Standort erfassten Beweggründe einbezogen. Für den Standort *Büro* wurden in 30 Interviews Angaben zu den Gründen der Standortwahl codiert; *tätigkeitsbezogene* Aspekte überwiegen dabei wie Tabelle 1 zu entnehmen ist. Unter den *tätigkeitsbezogenen* Gründen wurde das Bestreben Kolleg*innen zu treffen und sozial, u. a. auch informell und spontan beispielsweise in der Kaffeepause, zu interagieren angegeben. Dahinter können sowohl private Gründe (z.B. soziale Kontakte geschätzt, Einsamkeit) als auch eine berufliche Motivation liegen. Darüber hinaus kann die Wahl des Standort *Büro* in der Ausübung der Arbeitstätigkeit (und teils damit einhergehend einer Leistungssteigerung / Wahrung der Effizienz) begründet sein. Konkret fallen darunter die Nutzung eines angemessenen Arbeitsplatzes, der im *Homeoffice* möglicherweise nicht gegeben ist (z.B. Platz, Ausstattung, Sonderräume), die Vermeidung von Ablenkungen beispielsweise durch die Familie und die Wahl des *Büro* in Abhängigkeit der Arbeitsaufgabe. Wichtige, digital durchführbare Termine werden unter Umständen gezielt im *Büro* abgehalten. Auch Vorgaben oder eine Standort-Bindung bei Tätigkeiten können zur Wahl des Standort *Büro* führen. Bei Austausch-

Tätigkeiten (z.B. Arbeit im Team, Diskussion, Absprachen) wird ebenfalls teils das *Büro* gewählt. In einem Interview wird von besseren Ergebnissen durch persönliche Gespräche berichtet. Auch im Rahmen der erfassten Auswirkungen zeigt sich, dass in Folge der vermehrten digitalen Kommunikation, Präsenztermine gezielt abgehalten werden (n=21). Dazu zählen u. a. kreative Aufgaben, interaktive Tätigkeiten, Beziehungspflege, die Bearbeitung von Konflikten, Planungstätigkeiten und vertrauensvolle Gespräche. Darüber hinaus werden die weiteren Rahmenbedingungen des Termins wie beispielsweise Personenanzahl, Dauer und Wichtigkeit in die Entscheidung das Format betreffend einbezogen. In 21 Interviews wurde die Wichtigkeit der Organisation im Team angebracht. Mehrfach genannt wurden feste Anwesenheitstage, an denen das gesamte Team zusammen kommt, ggf. auch um gemeinsame Sitzungen abzuhalten.

Für die Wahl des *Homeoffice* wurden in allen 35 Interviews Motive codiert. Wie Tabelle 1 zu entnehmen ist, wurden *personenbezogene* Gründe ebenfalls in allen Interviews aufgeführt. *Tätigkeitsbezogene* Gründe wurden in 21 Interviews codiert. Die *tätigkeitsbezogenen* Gründe beziehen sich ausschließlich auf die Ausübung der Arbeitstätigkeit. In zwei Fällen wird allgemein von einer höheren Leistung bzw. Ergebnisqualität durch die Arbeit von zu Hause berichtet. Spezifische Gründe umfassen in wenigen Fällen eine bessere Ausstattung im *Homeoffice*, mehr Ruhe bzw. das Vermeiden von Ablenkungen (n=17) beispielsweise durch Kolleg*innen oder die tätigkeitsabhängige Wahl des *Homeoffice*. Letzteres bezieht sich beispielsweise auf Tätigkeiten, die viel Konzentration erfordern, allgemein konzeptionelle oder administrative Aufgaben, Lese- und Schreibarbeiten und besprechungsintensive Tage (digital).

Tabelle 1: Motive für die Nutzung des Büros und des Homeoffice (HO)

	Motiv	n Interviews
Büro	Personenbezogene Gründe	8
	Tätigkeitsbezogene Gründe, davon - Ausübung der Arbeitstätigkeit - Soziale Interaktion	30 (gesamt) 23 25
	Sonstiges	2
HO	Personenbezogene Gründe	35
	Tätigkeitsbezogene Gründe, davon - Ausübung der Arbeitstätigkeit	21 (gesamt) 21
	Sonstiges	7

Um psychischer bzw. körperlicher Belastung im *Homeoffice* zu begegnen, kann z. B. eine visuelle oder räumliche Abgrenzung des Arbeitsplatzes bzw. eine gute Ausstattung dienlich sein, wie im Rahmen der Interviews berichtet. Monopol des *Büro* bilden die Leistungserbringung als Team im Rahmen von Austausch-Tätigkeiten in Präsenz mit variierender Gruppengröße sowie die soziale Interaktion. Auch bei der Bewältigung einzelner Herausforderungen im Miteinander können gemäß der HOffEn-Auswertung z.B. Möglichkeiten der Begegnung oder Team-Events helfen. Die Nutzungsprozesse am Bürostandort zeichnen sich entsprechend durch eine stärkere Vielfalt aus verglichen mit der Einzelarbeit und den digitalen Austausch-Tätigkeiten im *Homeoffice*. In drei Interviews wurde der

Wandel des Büros hin zu einem Ort der Zusammenkunft angeregt. Spezifische, in den Interviews genannte Wünsche beinhalten Kreativareale, Workshop-Räume und Meetingtechnologien.

4 Diskussion

4.1 Ableitung von Bedarfen und Anforderungen an Arbeitsräumlichkeiten

In den folgenden Kapiteln 4.1.1 und 4.1.2 werden Bedarfe und Anforderungen an Arbeitsräumlichkeiten abgeleitet mit dem Bestreben eine insgesamt effiziente, nutzungsbezogene Flächenbereitstellung zu unterstützen. Personenbezogene Vorteile werden als Nebenprodukt/Anreiz für die Umsetzung neuer Arbeitskonzepte betrachtet. Grundannahme ist, dass die Arbeitsaufgaben unverändert bleiben und sich lediglich der Standort bzw. ggf. die Art und Weise der Bewältigung verändern.

4.1.1 Häuslicher Arbeitsplatz

Ziel ist die Gestaltung eines Arbeitsplatzes, der die Bewältigung der Arbeitsaufgaben unter Minimierung der psychischen und körperlichen Risiken ermöglicht. Der Arbeitsplatz *Homeoffice* wird im Folgenden als zusätzliche Wahlmöglichkeit betrachtet, die das Verrichten von Einzeltätigkeiten und digitaler Kommunikation mindestens im gleichen Maße wie der Büroarbeitsplatz unterstützen sollte. Die besondere oder gleichwertige Eignung der Arbeitsumgebung im *Homeoffice* kann, so unsere Schlussfolgerung, auf diese Weise gesamtheitlich betrachtet einen Mehrwert im Vergleich zum *Büro* darstellen (z.B. Verbesserung der individuellen Leistung oder Eliminierung von Störfaktoren aus dem Büroumfeld) und eine Anpassung der Flächennutzung am Standort *Büro* ermöglichen. Die nachfolgende Tabelle 2 zeigt die daraus folgenden Bedarfe und Anforderungen an den Homeoffice-Arbeitsplatz.

Tabelle 2: Übersicht der Bedarfe und Anforderungen an den Homeoffice-Arbeitsplatz

Nutzungsprozess	Bedarf	Anforderung
Ausübung der Arbeitstätigkeit	Körperliche Zuträglichkeit	Ergonomischer Arbeitsplatz
	Psychische Zuträglichkeit	Visuell, besser räumlich abgegrenzter Arbeitsbereich
	Störungsfreiheit	Bei Bedarf räumlich abgegrenzter Arbeitsbereich

Die Mindestanforderung gemäß Tabelle 2 ist die Unterbringung eines ergonomischen Arbeitsplatzes. Basierend auf den Vorgaben der DGUV [3] schließen wir auf einen Flächenbedarf von 2,46 m², besser 2,88 m². Ausreichend Wohnraum wurde auch im Zuge der Interviews als vorteilhaft eingestuft (n=6).

4.1.2 Büroarbeitsplatz

Die im *Büro* zu bedienenden Nutzungsprozesse sind die soziale Interaktion sowie die Ausübung der Arbeitstätigkeit. Wir schlussfolgern, dass interaktive Tätigkeiten (im Verhältnis zu Einzeltätigkeiten und rein digitaler Kommunikation) hier zunehmen, ggf. kumuliert an bestimmten Tagen stattfinden und gleichzeitig an Wichtigkeit gewinnen. Ziel ist folglich die Gestaltung eines Büroarbeitsplatzes, der den Schwerpunkt auf Austausch und die Arbeit im Team legt, zugleich aber den Bedarf an angemessenen Arbeitsplätzen für weitere Tätigkeiten deckt. Voraussetzung ist, dass grundsätzlich ausreichend Arbeitsplätze für alle Mitarbeitenden am Bürostandort zur Verfügung gestellt werden können, um die

Freiwilligkeit der Arbeit von zu Hause sowie die Umsetzung von Teamtagen zu gewährleisten. Tabelle 3 stellt die verschiedenen Tätigkeiten und die daraus resultierenden Anforderungen dar.

Tabelle 3: Übersicht der Bedarfe und Anforderungen an den Büroarbeitsplatz

Nutzungsprozess	Bedarf	Anforderung
Allgemein		Vielfältiges Raumangebot Arbeitsplätze für alle Mitarbeitenden
Soziale Interaktion	u. a. spontaner und informeller Austausch	Begegnungsfördernde Gebäudestrukturen mit zentralen Pausenbereichen u. Ä.
Ausübung der Abeitstätigkeit	Teamarbeit mit variierender Tätigkeit und Personenzahl / Diverse Einzelarbeiten und rein digitale Termine	Verschieden große Meetingräume z.B. Workshopräume, Kreativareale mit entsprechender Ausstattung / Reguläre, teils feste Arbeitsplätze in Einzel- und Gemeinschaftsbüros

In 20 Interviews wurde das Thema Desk-Sharing als zukünftig zu betrachtender Lösungsansatz aufgeführt. Ein Ansatz, um der erwarteten Tätigkeitsdiversität zu begegnen, stellt das Konzept des Activity Based Working (n=1) dar. Da bezogen auf das Teilen von Arbeitsplätzen verschiedene Bedenken geäußert wurden und eine Anwendung nicht immer möglich ist, sollten ggf. auch feste, persönliche Arbeitsplätze (n=1) eingerichtet werden. Allgemein ist zu beachten, dass bei geringer Kalkulierbarkeit zukünftiger Bedarfe (z.B. regelmäßige Personaländerungen) die Reduktion von Flächen im *Büro* nur schwer umsetzbar sein kann (n=1).

4.2 Diskussion der Methodik, Ergebnisse und Grenzen unserer Arbeit

Unsere Resultate stimmen teils mit anderen Studien über ein, die bestätigen, dass oftmals das *Homeoffice* für eine ruhige Arbeitsumgebung [9] und das *Büro* des Austauschs/Teams wegen geschätzt wird [10], [11], [9]. Auch verschiedene Tendenzen zur Flächenreduktion und Veränderungen der Flächennutzung (z.B. Desk-Sharing) wurden in weiteren Untersuchungen festgestellt [10], [11], [6]. In Anbetracht unserer Ergebnisse ist ungeklärt, ob/wie trotz des vielfältigeren Raumangebotes und weiterer Begrenzungen gesamtheitlich betrachtet Flächeneinsparungen durch effizientere Belegungskonzepte im *Büro* möglich sind. Verschiedene Studien zeigen, dass private Büros / Einzelbüros Vorteile mit sich bringen können [12], [13]. Ein Risiko bei Großraumbüros ist die Lärmbelästigung [14]. Auch ein Rückgang persönlicher Kommunikation zugunsten von digitaler Kommunikation ist möglich [15]. Tulenheimo-Eklund et al. [16] legen dar, dass unter Berücksichtigung der Bedürfnisse der Arbeitnehmenden ABW auch erfolgreich umgesetzt werden kann. Es stellt sich die Frage, ob bei Funktionstrennung gemäß Tätigkeit bzw. Konzentrationslevel die Einführung von Gemeinschaftsbüros vertretbar und gewinnbringend sein könnte. Die Arbeitsstättenrichtlinie [17] gibt vor, dass Räume mit einem Arbeitsplatz mindestens 8 m² groß sein müssen; für jeden weiteren Arbeitsplatz sind mindestens 6 m² gefordert. Welches Einsparpotential sich praktisch ergibt ist genauer zu untersuchen. Sollte eine Erweiterung der Flächen, die für Arbeiten im Team zur Verfügung stehen erforderlich sein, könnten Räume, die eine kurzfristige Umfunktionalisierung zu anderweitigen Arbeitsplätzen zulassen, möglicherweise Auslastungsspitzen abdecken. Ob Kommunikation nach wie vor ein Problem darstellt,

ist basierend auf dem Ergebnisreports der Konstanzer Homeoffice Studie [2] kritisch zu hinterfragen - ggf. sind die Relevanz und Signifikanz der erarbeiteten Anforderungen zu erörtern.

Neben der Flächenmenge wird der effiziente Betrieb von Räumlichkeiten auch durch Energieverbräuche beeinflusst. Eine selektive Temperierung von genutzten Flächen durch entsprechende Gliederung könnte vorteilhaft sein. Der Vergleich von Sensordaten der beiden Standorte könnte zudem Aufschluss darüber geben, wo Anforderungen unter- oder überschritten werden.

Eine Beeinflussung der Ergebnisse durch Nachwirkungen der Pandemie kann nicht vollständig ausgeschlossen werden. Allgemein lässt sich schlussfolgern, dass Vorteile hinsichtlich der Leistungsfähigkeit an beiden Standorten gegeben sein können. Es ist anzumerken, dass die Motive einzelner Personen von der individuellen Bewertung abhängen. Eine Arbeitsumgebung, die Person A an der Bewältigung ihrer Arbeitstätigkeit hindert, kann Person B durchaus als angemessen bewerten. Neben der kleinen Stichprobengröße können auch durch die Eigenheiten der Methodik Verzerrungen entstehen, die die Aussagekraft quantitativer Beobachtungen schwächen.

5 Fazit & Ausblick

Umnutzungen von Büroflächen sollten unter Berücksichtigung der Tätigkeiten sowie der Bedürfnisse der Mitarbeitenden abgewogen werden. Die durch den Wechsel zu effizienteren Belegungskonzepten vermuteten Flächeneinsparungen im *Büro* scheinen grundsätzlich möglich. Inwiefern dadurch steigende Flächenbedarfe im Privatbereich ausgeglichen werden können, ist mittels differenzierter Betrachtung und Diskussion räumlicher Bedarfe zu untersuchen. Ebenso scheint eine Standort-übergreifende Betrachtung energetischer Auswirkungen zielführend. Weitere Untersuchungen bezüglich nachhaltiger und sozialer Planung der räumlichen Umsetzung neuer Arbeitskonzepte sind vorgesehen.

Acknowledgements

Wir erkennen die Förderung durch das Projekt „HOffEn - Home-Office und Energiewende“ (BMWK (BMWE), Förderkennzeichen 03EI5228) dankend an, welche uns die Umsetzung der wesentlichen Teile dieser Arbeit ermöglicht hat. Wir danken unseren Projektpartner*innen für die Zusammenarbeit.

Literatur

[1] ifo Institut, *Homeoffice-Quote stabilisiert sich bei knapp 25 %*, https://www.ifo.de/fakten/2025-03-24/homeoffice-quote-stabilisiert-sich-bei-knapp-25-prozent, 2025.

[2] F. Kunze und K. Hampel, *Zwischen Präsenzpflicht und Homeoffice-Euphorie : Stand des mobilen Arbeitens fünf Jahre nach Beginn der Corona-Pandemie : Konstanzer Homeoffice Studie Ergebnisreport April 2025*, http://nbn-resolving.de/urn:nbn:de:bsz:352-2-rrsuie9779vv4, 2025.

[3] DGUV, Hrsg., *DGUV Information 215-441 Büroraumplanung*, https://publikationen.dguv.de/regelwerk/dguv-informationen/224/bueroraumplanung-hilfen-fuer-das-systematische-planen-und-gestalten-von-bueros?c=15, Berlin, 2023.

[4] S. Herwegen, G. Flesch und R. Stelten, »6.1 Wohnen«, in *Sozial Bericht 2024*. Bonn, 2024. Adresse: https://www.bpb.de/kurz-knapp/zahlen-und-fakten/sozialbericht-2024/

[5] S. Krause, A. Trumpp, T. Dichtl, S. Kiese und A. Rutsch, »Neue Arbeitswelt, neue Arbeitsorte: Auswirkungen von Homeoffice auf den Büroimmobilienmarkt«, *ifo Schnelldienst*, Jg. 77, Nr. 03, S. 63–73, 2024. Adresse: https://hdl.handle.net/10419/296213

[6] S. Krause, A. Trumpp, T. Dichtl, S. Kiese und A. Rutsch, »Homeoffice und die Zukunft der Büros: Flexibilisierung, Reduzierung und Umnutzungspotenzial«, *ifo Schnelldienst*, Jg. 77, Nr. 09, S. 64–68, 2024. Adresse: https://hdl.handle.net/10419/306130

[7] P. Mayring, *Einführung in die qualitative Sozialforschung*, 7. Aufl. Beltz Verlagsgruppe, 2023.

[8] A. Aretz u. a., »Home-Office und Energiewende (HOffEn): Veränderungen durch mobiles Arbeiten und Chancen für die Transformation des Energiesystems«, 2025. DOI: 10.34657/19255

[9] IBA, Hrsg., *Status quo der Büroarbeit in Deutschland. Update*, https://iba.online/site/assets/files/16432/iba-studie-2023-24.pdf, 2024.

[10] Drees & Sommer SE, *Trendstudie Workspace Benchmark*, https://www.dreso.com/de/workspace-benchmark-report#kontaktformular, 2023.

[11] Gensler Research Institute, *Die Rückkehr ins Büro*, https://www.gensler.com/gri/germany-workplace-survey-2023-de, 2023.

[12] J. Radun und V. Hongisto, »Perceived fit of different office activities – The contribution of office type and indoor environment«, *Journal of Environmental Psychology*, Jg. 89, 2023. DOI: 10.1016/j.jenvp.2023.102063

[13] A. Gerlitz und M. Hülsbeck, »The productivity tax of new office concepts: a comparative review of open-plan offices, activity-based working, and single-office concepts«, *Management Review Quarterly*, Jg. 74, Nr. 2, S. 745–775, 2024. DOI: 10.1007/s11301-022-00316-2

[14] O. James, P. Delfabbro und D. L. King, »A Comparison of Psychological and Work Outcomes in Open-Plan and Cellular Office Designs: A Systematic Review«, *SAGE Open*, Jg. 11, Nr. 1, 2021. DOI: 10.1177/2158244020988869

[15] E. S. Bernstein und S. Turban, »The impact of the 'open' workspace on human collaboration«, *Philosophical Transactions of the Royal Society B: Biological Sciences*, Jg. 373, Nr. 1753, 2018. DOI: 10.1098/rstb.2017.0239

[16] E. Tulenheimo-Eklund, A. Haapakangas, M. Hirvonen, V. Ruohomäki und K. Reijula, »Perceptions of activity-based offices are associated with employee well-being and self-reported work ability in hybrid work: a cross-sectional study«, *Journal of Occupational Health*, Jg. 67, Nr. 1, 2025. DOI: 10.1093/joccuh/uiaf027

[17] ASTA, *ASR A1.2 Raumabmessungen und Bewegungsflächen*, Berlin, 2013/2022. Adresse: https://www.baua.de/DE/Angebote/Regelwerk/ASR/ASR-A1-2

Utilization of deep autoencoders for improving semantic segmentation of 3D point clouds

Marcel Aubart and Hristo Vassilev
Chair of Computing in Civil Engineering and Geoinformation Systems and Geodetic Institute (gia),
RWTH Aachen University, Mies-van-der-Rohe-Str. 1, 52074 Aachen, Germany
E-mail(s): marcel.aubart@gia.rwth-aachen.de, hristo.vassilev@gia.rwth-aachen.de

Abstract: Machine learning is increasingly explored to create semantically rich three-dimensional as-is models (as-is Building Information Modeling (BIM)) of infrastructure assets, which form the basis of digital twins. Point clouds captured from these assets are segmented and classified to identify distinct objects or building components, allowing the generation of three-dimensional (3D) models. However, challenges such as occlusion, low resolution, scan artifacts, and noise may hinder performance, especially when labeled training data are limited. These problems can be mitigated using improved scanners, increasing scan positions, applying rule-based techniques, or employing generative machine learning models. This paper presents a pre-training method for bridge structure point clouds that leverages unlabeled data to enhance semantic segmentation. The approach involves three steps. First, realistic occlusions are simulated by masking points that are not visible from a semi-random camera perspective. Next, an encoder-decoder network is trained to reconstruct the occluded areas, thereby learning spatial structures. Finally, the encoders pre-trained weights are used to initialize a semantic segmentation network. Experiments on a Mobile Laser Scanning dataset show improvements in accuracy and mean Intersection over Union for architectures including PointNet, Point Completion Network (PCN), and Dynamic Graph Convolutional Neural Network (DGCNN).

Keywords: Building Information Modeling (BIM), Machine Learning, Point Cloud Segmentation

DOI: 10.18154/RWTH-CONV-254912. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

The rapid advancement in reality capture has revolutionized civil engineering by providing a data basis for the automated generation of semantically rich as-is models (as-is BIM) for infrastructure assets. Such detailed models are essential for monitoring the condition of structures that naturally deteriorate over time, such as bridges and other critical constructions.

Despite these developments, significant challenges remain. Semantic segmentation of point clouds, the task of assigning meaningful categories to individual points, is particularly demanding. Even with modern deep learning techniques, issues such as artifacts, low resolution, noise, and occlusions impede reliable segmentation. Occlusions, which occur inherently when parts of a structure are hidden by perspective or sensor limitations, pose a severe hurdle. Additionally, the scarcity of labeled training

data further complicates the development of robust supervised segmentation models, while unlabeled data remain largely unused.

To address these challenges, this paper proposes an unsupervised pre-training approach that takes advantage of occlusion networks, a subset of generative models, to improve the quality of semantic segmentation. Central to the method is an encoder–decoder framework trained on artificially occluded data that closely resembles the naturally occurring occlusions in bridge structures. By pre-training the encoder on an occlusion completion task, the approach extracts general features from extensive unannotated point clouds. Subsequently, the encoder is fine-tuned within a semantic segmentation framework using a limited set of annotated data. This two-step process aims to improve the accuracy, robustness, and overall generalization of the segmentation model.

2 State-of-the-Art

This chapter gives an overview of the challenges of converting raw point cloud scans into as-is models (as-is BIM), such as occlusions. It evaluates topics such as unsupervised pre-training, chamfer distance, and occlusion simulation to enhance semantic segmentation accuracy.

2.1 Scan to BIM

Scan-to-BIM refers to the conversion of raw point cloud scans into as-is models (as-is BIM) including automatic processes like semantic segmentation, where each point is labeled to identify structural elements (e.g., beams, columns, decks). However, inherent challenges in point cloud data can hinder segmentation quality. One major issue are occlusions, resulting from physical obstructions or sensor limitations that hide parts of the structure, resulting in incomplete or fragmented representations. In addition, low resolution and noise can blur fine details and distort features, leading to misclassifications that affect the overall integrity of the BIM [1]. Errors in segmentation can propagate through the entire Scan-to-BIM process, undermining its accuracy and reducing its utility for tasks such as condition monitoring and digital twin development.

To mitigate these challenges, current research incorporates techniques such as rule-based preprocessing to standardize and enhance raw scan data, and unsupervised pre-training methods, such as occlusion completion [2].

2.2 Unsupervised Pre-training

Unsupervised pre-training has become increasingly popular due to its demonstrated success in various domains, including natural language understanding [3] and object detection [4]. In the case of point clouds, the two predominant unsupervised pre-training strategies are generative modeling and self-supervised learning. Generative modeling approaches for point clouds have included, for example, generative adversarial networks (GANs) [5] and autoencoders [6]. Despite these efforts, recent work has shown that self-supervised methods outperform generative models in this context [7].

2.3 Chamfer Distance

The Chamfer distance [8] is a widely used loss function in point cloud reconstruction. It computes the average distance from each point in one set to its nearest neighbor in the other, making it efficient and robust to variations in point density. Minimizing this distance allows the encoder–decoder model to

capture both global structures and fine details. Since it relies only on geometric properties, it is well suited for unsupervised learning.

2.4 Occlusion Simulation

Occlusion simulation is critical for generating realistic training data for point cloud completion. In practice, virtual cameras are deployed to simulate different sensor viewpoints, and ray-casting techniques are deployed to determine visible and occluded regions. This process replicates real-world scanning conditions in which parts of a structure are hidden because of perspective or sensor limitations. By introducing controlled occlusions, reconstruction models are challenged to infer missing information, thereby learning more robust spatial representations. These enhanced representations not only improve point cloud completion, but also contribute to better performance in downstream tasks such as semantic segmentation and digital reconstruction.

3 Methodology

The overall approach (Fig.: 1) involves generating occluded point clouds from multiple virtual camera viewpoints, training an occlusion completion model to reconstruct complete point clouds, and then initializing downstream semantic segmentation models with and without the pre-trained encoder weights for a comparison, the latter one will be used as a baseline.

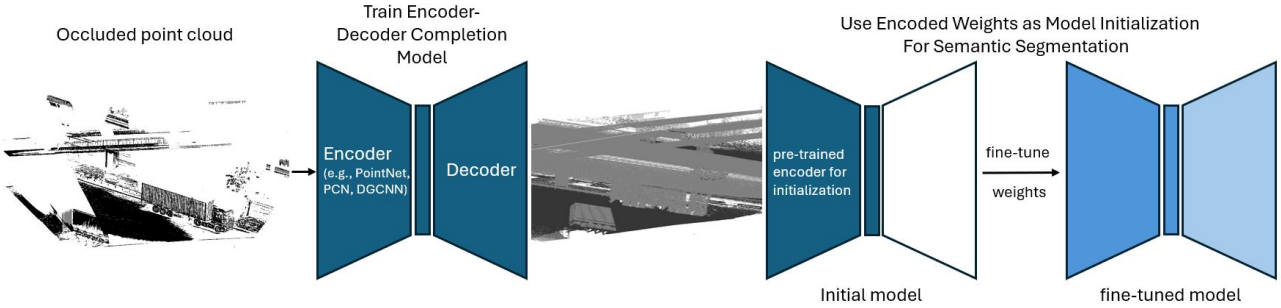


Figure 1: From left to right the generated occluded point cloud data, which is trained in an encode-decoder completion model. Afterwards, the pre-trained encoder is utilized in the initial model of the semantic segmentation. Fine-tuning the weights, results in a fine-tuned semantic segmentation model.

3.1 Data Preprocessing of the unlabeled data

Before occlusion generation, unlabeled point clouds are preprocessed to ensure consistency and readiness for further processing. First, a rule-based algorithm extracts the XYZ coordinates and assigns a default class to each point. Points belonging to the class ground are identified using the lowest 35% of the Z values, the mean of these points is assumed to be the zero level. The difference between the old mean and the new zero level is then subtracted from all Z-values of the point cloud. Points above a 3.5m threshold are analyzed via a histogram to detect dense layers corresponding to the bridge deck, these are classified as bridge points, while all remaining points are labeled as other structures.

In a next step, the point cloud is rotated. The bridge points are centered and processed using principal component analysis (PCA) to determine the primary axis of the structure. The angle between the first principal component and the X-axis is used to rotate the point cloud so that the bridge aligns with the X-axis, after which the cloud is recentered in the XY plane.

Finally, to simulate occlusions, point clouds are converted into meshes. This conversion involves optional voxel-based down-sampling to reduce computational load while preserving geometric details, followed by normal estimation and reorientation to the local tangent plane. Poisson surface reconstruction then generates a three-dimensional (3D) mesh, and low-density vertices are filtered out before the mesh is clipped to the original bounding box. This preprocessing pipeline produces uniformly aligned point clouds and high-quality meshes, ready for subsequent occlusion simulation.

3.2 Generating Occlusions

In this section, the algorithm used for occlusion generation is presented. The method simulates sensor occlusions by artificially occluding point clouds from various virtual camera viewpoints and computing point visibility with ray-casting techniques. This process produces data that closely mimic the occlusions encountered in real-world sensor data. The input to this algorithm consists of preprocessed, aligned point clouds representing the complete structure along with their preprocessed corresponding mesh representations.

3.2.1 Virtual Camera Positioning

To simulate sensor perspectives, virtual cameras are positioned around the bridge. They are placed along the Y-axis with varying X-coordinates to surround the scene from different lateral angles. The number of cameras and their elevations are adjustable, with cameras divided into two groups on opposite sides of the structure. Their horizontal placement is determined by scaling the point cloud's extent and applying appropriate padding.

Each camera is oriented toward a fixed target point near the bridge center for a consistent viewpoint. To establish each camera's coordinate frame, the forward direction (from the camera to the target) is computed first. Then, a right vector is derived by taking the perpendicular to the forward direction and the global upward axis, and finally, an up vector is obtained as perpendicular to both. These three orthogonal directions form the basis for the camera's rotation. When combined with the cameras position, they create the extrinsic transformation matrix used for transforming global coordinates.

In addition, intrinsic parameters such as focal length are defined to control the field of view and projection scale, ensuring that the virtual cameras accurately simulate real-world imaging.

3.2.2 Visibility Computation Using Ray Casting

Once the virtual cameras are positioned, visible points are determined via ray casting. To reduce computation while preserving geometry, the full point cloud is down-sampled using voxel-based aggregation, and mesh normals are computed and oriented outward.

For each camera, rays are cast from its position toward the down-sampled points, and only those falling within the camera's field of view (FoV) (selected by comparing the ray direction with the camera's forward vector) are considered. The distance to the first intersection (hit) is compared to the Euclidean distance from the camera to each point: if the hit distance is greater than or equal to the ray length (or infinite), the point is deemed visible. Otherwise, it is occluded. The visible points are then aggregated to form the occluded point cloud for that camera. Afterwards, each cloud is saved in multiple overlapping blocks, using a naming convention that includes the original point clouds identifier, the location in the original cloud and the camera index. This enables further processing and analysis in the occlusion completion task.

3.3 Occlusion Completion

The point cloud completion model reconstructs a point cloud from a sparse, occluded input using a coarse-to-fine strategy. The process has three parts: a coarse generator, a fine generator, and a feature extractor. The process is evaluated on three encoders PointNet [9], PCN [10] and DGCNN [11] within the same framework. Each encoder encodes an occluded point cloud into a 1024-dimensional latent vector. For reconstruction, the folding-based decoder from PCN [10] is modified to reconstruct blocks of the occluded point cloud in two stages. First, the decoder produces a coarse reconstruction with 512 points, which is refined by warping a 4×4 grid around each point to enable a fine reconstruction of 8192 points. The model is optimized using the Chamfer distance [8] loss function (Section 2.3).

4 Experimental Setup

This section describes the experimental setup used to evaluate the proposed decoder-encoder approach for improving the semantic segmentation task, including details about the model architecture and the training protocol.

4.1 Dataset

The dataset used, comprises 384 highway scenes captured by a Mobile Laser Scanning (MLS) system in North Rhine-Westphalia, Germany, featuring two kinds of bridges: sign gantries and concrete bridges. It includes 69 manually segmented point clouds (with 11 merged classes from an initial 18) and 315 unlabeled ones. Each class represents objects, which are recognizable from the point cloud geometry alone, since only geometry data and intensity features were available. The occluded point clouds are generated with 12 virtual cameras (Section 3.2.1).

4.2 Training of the Completion Task

As mentioned before, the pre-training completion model is an encoder-decoder model. To demonstrate that the pre-training method is independent of the network architecture, three different encoders PointNet [9], PCN [10], and DGCNN [11] are employed. Each encoder converts the occluded point clouds into a latent vector, while a modified folding-based decoder reconstructs the complete shape in two stages. This enables us to assess the impact of different encoders under identical conditions.

4.3 Training of the Semantic Segmentation Task

The semantic segmentation task partitions a point cloud into distinct regions through point-level classification. An encoder extracts hierarchical, spatial features from the point cloud, and a decoder projects these features back to 3D space to assign class labels. To evaluate the benefits of pre-training, we compare segmentation models initialized with pre-trained encoder weights against those with random initialization.

4.4 Training Protocol

Both the pre-training and segmentation models are optimized using AdamW [12]. For pre-training, a learning rate of 1×10^4 and a batch size of 32 is used. The model was trained for 50 epochs.
For the semantic segmentation task, a decaying learning rate (initially 1×10^4, decaying after 20 steps) is applied. The PointNet [9] and PCN [10] models were trained for 800 epochs, while DGCNN [11] models were trained for 600 epochs with 20 nearest neighbors for graph construction. The batch size during semantic segmentation training was set to 8 to accommodate higher memory demands.

5 Results

This section presents the results of the experiments for the semantic segmentation task of bridge structures with and without utilization of a pre-trained encoder on artificial generated occluded point cloud data.

Table 1 presents the performance of the six models-PointNet [9] baseline, PointNet [9] with a pre-trained encoder, PCN [10] baseline, PCN [10] with a pre-trained encoder, DGCNN [11] baseline, and DGCNN [11] with a pre-trained encoder. Each model was evaluated on overall accuracy, mean Intersection over Union (mIoU), and class-specific Intersection over Union (IoU) values across eleven classes. The table also highlights the performance improvements achieved by replacing random encoder weights with a pre-trained encoder, thereby illustrating the benefits of occlusion completion pre-training.

Table 1: Combined performance metrics for PointNet [9], PCN [10], and DGCNN [11] with baseline and occlusion completion pre-trained encoder initialization, along with the corresponding performance differences.

Model	Condition	Accuracy %	mIoU %	unlabeled IoU%	assembly_abutment IoU%	assembly_signbridge IoU%	barrier IoU%	beam IoU%	column_footing IoU%	natural_ground IoU%	slab_deck IoU%	slab_paving IoU%	traffic_sign IoU%	wall_noisebarrier IoU%
PointNet [9]	Baseline	81.0	40.6	9.2	31.4	40.7	41.1	47.5	15.3	55.3	57.2	90.6	42.6	15.3
	Pre-trained	83.6	43.7	5.1	36.5	47.5	42.7	54.5	20.3	63.7	60.1	92.6	44.7	12.8
	Δ	2.6	3.1	−4.1	5.1	6.8	1.6	7.0	5.0	8.4	3.5	2.0	2.1	−2.5
PCN [10]	Baseline	76.6	31.8	10.1	28.4	27.9	40.4	24.8	7.5	47.8	47.0	88.5	23.2	3.7
	Pre-trained	79.6	37.1	4.4	30.3	46.9	33.9	41.7	9.8	57.0	49.7	90.1	34.3	9.8
	Δ	3.0	5.3	−5.7	1.9	19.0	−6.5	16.9	2.3	10.0	2.7	1.6	11.1	5.3
DGCNN [11]	Baseline	88.3	53.2	7.9	49.5	47.2	52.3	72.7	18.2	71.9	75.4	94.1	58.5	25.4
	Pre-trained	89.0	54.8	14.4	52.7	60.3	52.4	74.3	20.7	73.5	77.8	95.8	53.0	27.5
	Δ	0.7	1.6	6.5	3.2	13.1	0.1	1.6	2.5	1.6	2.4	1.7	−5.5	2.1

6 Discussion

The results in Table 1 demonstrate that using a pre-trained encoder for occlusion completion improves semantic segmentation across all models. Accuracy increased by 2.6% for PointNet [9], 3.0% for PCN [10], and 0.7% for DGCNN [11], while mean Intersection over Union improved by 3.1%, 5.3%, and 1.6%, respectively. Classes closely connected to the structure, such as $slab_deck$, $assembly_signbridge$, and $beam$ showed notable gains, indicating that pre-training helps especially in regions prone to occlusion. Although minority classes (e.g., $column_footing$) also improved, they remained challenging due to limited data and inherent occlusions. In contrast, classes further from the structure experienced slight declines, likely caused by the targeted generation of occlusions around the bridge. Pre-training not only enhanced performance but also sped up convergence during semantic segmentation training, thereby reducing the number of epochs required. However, the occlusion completion pre-training

increases memory demands, especially for DGCNN [11], resulting in smaller batch sizes and slower training. Better hardware and the ability to tune hyperparameters extensively could further improve these results. Furthermore, while PCN [10] showed the highest relative improvement with pre-training, its overall performance remained below that of PointNet [9] and DGCNN [11] due to its original design focus on completion rather than segmentation.

7 Conclusion and Outlook

This paper demonstrates that incorporating a pre-trained encoder for occlusion completion significantly improves semantic segmentation of bridge structures. The results show consistent gains in accuracy and IoU across PointNet [9], PCN [10], and DGCNN [11], especially for classes representing structural elements. Pre-training helps models learn richer shape representations, making them more robust to incomplete data, which is crucial for datasets plagued by occlusions, as is common in bridge scans. Additionally, pre-training accelerates convergence by reducing the number of training epochs required. Initializing the models with weights learned from occlusion completion provides prior shape knowledge that streamlines fine-tuning, an important advantage when computational resources are limited. While hardware constraints limited batch sizes and grid resolutions during occlusion completion, potentially capping performance gains, more powerful hardware could allow for larger configurations and further enhance segmentation performance. Although PCN [10] experienced the highest relative improvement from pre-training, its overall performance remained lower than that of PointNet [9] and DGCNN [11], likely due to its original design focus on completion rather than segmentation. Future work could explore more advanced segmentation architectures (e.g., PointNeXt) and alternative occlusion strategies, as well as assess the impact of higher quality point cloud data (e.g., from Terrestrial Laser Scanners (TLS)) on segmentation accuracy. The enhanced features from occlusion completion may also benefit other tasks, such as geometry fitting using RANSAC to reconstruct object boundaries from incomplete data. Overall, these findings underscore the premise of unsupervised occlusion completion pre-training for robust point cloud processing in challenging environments.

References

[1] J. Martens, T. Blut, and J. Blankenbach, “Cross domain matching for semantic point cloud segmentation based on image segmentation and geometric reasoning”, *Advanced Engineering Informatics*, vol. 57, p. 102 076, 2023. DOI: 10.1016/j.aei.2023.102076

[2] H. Wang, Q. Liu, X. Yue, J. Lasenby, and M. J. Kusner, “Unsupervised point cloud pre-training via occlusion completion”, in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021, pp. 9782–9792. DOI: 10.48550/arXiv.2010.01089

[3] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding”, in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1, 2019, pp. 4171–4186. DOI: 10.18653/v1/N19-1423

[4] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, “Momentum contrast for unsupervised visual representation learning”, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. DOI: 10.48550/arXiv.1911.05722

[5] P. Achlioptas, O. Diamanti, I. Mitliagkas, and L. Guibas, “Learning representations and generative models for 3D point clouds”, in *Proceedings of the 35th International Conference on Machine Learning*, vol. 80, 2018, pp. 40–49. DOI: 10.48550/arXiv.1707.02392

[6] Y. Yang, C. Feng, Y. Shen, and D. Tian, “FoldingNet: Point cloud auto-encoder via deep grid deformation”, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018. DOI: 10.48550/arXiv.1712.07262

[7] S. Xie, J. Gu, D. Guo, C. R. Qi, L. Guibas, and O. Litany, “PointContrast: Unsupervised pre-training for 3d point cloud understanding”, in *Computer Vision - ECCV 2020*, Springer International Publishing, 2020, pp. 574–591. DOI: 10.1007/978-3-030-58580-8_34

[8] M. Akmal Butt and P. Maragos, “Optimum design of chamfer distance transforms”, *IEEE Transactions on Image Processing*, vol. 7, no. 10, pp. 1477–1484, 1998. DOI: 10.1109/83.718487

[9] R. Q. Charles, H. Su, M. Kaichun, and L. J. Guibas, “Pointnet: Deep learning on point sets for 3d classification and segmentation”, in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 77–85. DOI: 10.1109/CVPR.2017.16

[10] W. Yuan, T. Khot, D. Held, C. Mertz, and M. Hebert, “PCN: Point completion network”, in *2018 International Conference on 3D Vision (3DV)*, 2018, pp. 728–737. DOI: 10.1109/3DV.2018.00088

[11] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon, “Dynamic graph cnn for learning on point clouds”, *ACM Transactions on Graphics*, vol. 38, no. 5, Oct. 2019. DOI: 10.1145/3326362

[12] I. Loshchilov and F. Hutter, “Decoupled weight decay regularization”, in *7th International Conference on Learning Representations, ICLR 2019*, 2019. DOI: 10.48550/arXiv.1711.05101

Untersuchung zur Akzeptanz und Nutzung von Energie- und Komfortmonitoringsystemen in privaten Haushalten

Lisa von Wittenhorst zu Sonsfeld[1] and Elisabeth Beusker[1]
[1]Lehr- und Forschungsgebiet für Immobilienprojektentwicklung (iPE), RWTH Aachen University, Aachen, Germany
E-mail(s): lisa.vonwittenhorst@ipe.arch.rwth-aachen.de, elisabeth.beusker@ipe.arch.rwth-aachen.de

Abstract: Die Energiewende erfordert neben technologischen Innovationen auch eine aktive Verhaltensänderung der Gesellschaft im Umgang mit Energie. Digitale Monitoringsysteme können diesen Wandel unterstützen, indem sie Verbrauchstransparenz schaffen, Einsparpotentiale sichtbar machen und Verhaltensanpassungen anstoßen. Voraussetzung für ihre Wirksamkeit ist jedoch eine breite gesellschaftliche Akzeptanz und Nutzung. Im Rahmen der vorliegenden Arbeit wird daher der Einfluss soziodemografischer und wohnstruktureller Faktoren auf die Akzeptanz und Nutzung sowie die Unterschiede in den Nutzungsmotiven und -hemmnissen ermittelt. Grundlage der Untersuchung bildet eine standardisierte Online-Befragung von etwa 280 privaten Haushalten, die im Zeitraum von 2022 bis 2023 im Rahmen des Reallabors der Energiewende *SmartQuart* durchgeführt wurde. Die Datenanalyse erfolgt mittels Mittelwertvergleichen und binär-logistischer Regressionsverfahren. Die Ergebnisse zeigen, dass funktionale Motive wie Kostenreduktion und Energieeinsparung zentrale Treiber der Nutzung von Energiemonitoring darstellen, während beim Komfortmonitoring das Motiv der Wohnkomfortsteigerung dominiert. Hohe Anschaffungskosten wirken sich bei beiden Technologien nutzungshemmend aus. Beim Komfortmonitoring wurde zudem das fehlende Interesse als Barriere identifiziert. Technische Hürden, Datenschutzbedenken und Informationsdefizite spielen hingegen eine untergeordnete Rolle. Die Nutzung der Systeme hängt darüber hinaus signifikant von Faktoren wie Bildungsniveau, Einkommenshöhe, Wohnfläche und Gebäudetyp ab. Die Ergebnisse bieten differenzierte Einblicke in akzeptanzrelevante Faktoren von Monitoringsystemen, sind aufgrund einer nicht-repräsentativen Stichprobe jedoch nur eingeschränkt generalisierbar.

Keywords: Nutzerverhalten, Energiemonitoring, Komfortmonitoring, Akzeptanz, Private Haushalte

DOI: 10.18154/RWTH-CONV-254902. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Einführung

Im Kontext der Energiewende spielt die Reduktion des Energieverbrauchs im Gebäudesektor eine zentrale Rolle. Wohngebäude sind in Deutschland für einen signifikanten Anteil des Endenergieverbrauchs verantwortlich, insbesondere durch Raumwärme und Warmwasserbereitstellung [1]. Um den

Energiebedarf in diesem Bereich wirksam zu senken, sind neben technischen Maßnahmen wie energetischer Sanierung oder dem Einsatz erneuerbarer Energiequellen auch Verhaltensänderungen der Nutzer[1] erforderlich. Verhaltensbezogene Maßnahmen bieten ein großes, bislang nur unzureichend ausgeschöpftes Potential zur Energieeinsparung. Studien zeigen, dass durch bewusstes Nutzerverhalten – etwa durch optimiertes Heiz- und Lüftungsverhalten, die Reduktion von Standby-Verlusten oder die Anpassung der Raumtemperatur – erhebliche Einsparungen erzielt werden können [2, 3, 4]. Die Umsetzung dieser Maßnahmen erweist sich im Alltag allerdings als herausfordernd, da sie ein hohes Maß an Aufmerksamkeit, Motivation und Wissen voraussetzen.

An dieser Stelle setzen digitale Monitoringsysteme an, indem sie Verbrauchswerte in Echtzeit detailliert erfassen und visualisieren. Sie schaffen damit Transparenz, fördern ein bewussteres Verhalten und können die Bildung energieeffizienter Gewohnheiten unterstützen [5, 6, 7]. Die Integration von Monitoringsystemen in privaten Haushalten ist ein ergänzendes Instrument zur technischen Modernisierung, da sie verhaltensbasierte Energieeinsparungen unterstützt. Im Gegensatz zu steuernden Systemen dienen Monitoringsysteme ausschließlich der Erfassung und Analyse von Verbrauchs- und Komfortdaten. Im Rahmen dieser Arbeit werden zwei Typen von Systemen betrachtet: Energiemonitoring und (Nutzer-)Komfortmonitoring. Energiemonitoring umfasst Energieverbräuche, beispielsweise von Strom, Wärme oder Wasser. Die Erfassung der Verbräuche ermöglicht es, Einsparpotentiale zu identifizieren und die Nutzung technischer Anlagen zu optimieren. Beim Komfortmonitoring stehen dagegen raumklimatische Parameter wie Temperatur, Luftfeuchtigkeit und CO_2-Konzentration im Fokus. Die Analyse dieser Daten dient der Optimierung des Raumklimas sowie der Steigerung der Nutzerzufriedenheit und der Produktivität. In der Regel basieren beide Systemtypen auf funkgestützten Infrastrukturen, wie digitalen Zählern und Sensoren, die über zentrale Plattformen organisiert sind und sich auch nachträglich in Gebäude integrieren lassen. Durch die Kombination von Energie- und Komfortmonitoring und durch die gezielte Anpassung des Nutzerverhaltens ist es möglich, eine gleichzeitige Steigerung von Energieeffizienz und Wohnkomfort bei minimalem Ressourceneinsatz zu erreichen. Ihre Wirksamkeit hängt jedoch maßgeblich von ihrer gesellschaftlichen Akzeptanz und flächendeckenden Nutzung ab.

Im Rahmen dieser Arbeit werden daher Einflussfaktoren auf die Akzeptanz und Nutzung von Monitoringsystemen in privaten Haushalten untersucht. Im Zentrum der Untersuchung stehen sowohl subjektiv bewertete Nutzungsmotive und Hemmnisse als auch soziodemografische und wohnstrukturelle Einflussfaktoren. Das Ziel der Arbeit ist es, anhand der ermittelten Einflussfaktoren ein tieferes Verständnis für die zielgruppenspezifischen Beweggründe und Hürden bei der Nutzung der verschiedenen Monitoringsysteme zu gewinnen und damit einen Beitrag zur digitalen Transformation privater Haushalte zu leisten.

2 Methode

Die Untersuchung basiert auf einer Befragung, die im Rahmen des vom Bundesministerium für Wirtschaft und Energie (BMWE) geförderten Reallabors der Energiewende *SmartQuart* durchgeführt wurde. Seit 2020 untersucht das Projekt die Umsetzung der Energiewende in drei deutschen Quartieren unter realen Bedingungen. Dabei werden sozioökonomische und technische Herausforderungen

[1] Aus Gründen der besseren Lesbarkeit wird auf die gleichzeitige Verwendung mehrerer Sprachformen verzichtet. Sämtliche Personenbezeichnungen gelten für Personen jeglichen Geschlechts.

für unterschiedliche Akteursgruppen systematisch identifiziert und analysiert. Zur empirischen Erhebung wurde auf Basis einer umfassenden Literaturrecherche ein standardisierter Online-Fragebogen entwickelt, der mit dem Tool SoSci Survey umgesetzt wurde [8], [9]. In dieser Arbeit werden ausgewählte Ergebnisse präsentiert, die sich auf die Nutzung digitaler Monitoringsysteme im privaten Wohnumfeld beziehen.

Die erhobenen Daten wurden mithilfe statistischer Verfahren ausgewertet, um differenzierte Erkenntnisse über die Einflussfaktoren auf die Nutzung und Implementierung digitaler Monitoringsysteme zu erlangen. Die Untersuchung wurde in zwei Schritten durchgeführt: Zunächst wurden die Motive und Hemmnisse für die Implementierung der verschiedenen Systeme erfasst und mittels eines Mittelwertvergleichs ausgewertet. Im zweiten Schritt wurde mittels einer binär-logistischen Regression der Einfluss soziodemografischer und wohnstruktureller Faktoren auf die Nutzung untersucht.

3 Ergebnisse

3.1 Untersuchung der Stichprobe

Die Untersuchungen basieren auf den Angaben von insgesamt 284 privaten Nutzern, die im Zeitraum von September 2022 bis Dezember 2023 an der Befragung teilgenommen haben. Die Stichprobe weist einen hohen Männeranteil auf: 74,5% (n=205) der Teilnehmer sind männlich, während 25,5% (n=70) weiblich sind. Die Altersverteilung zeigt, dass knapp ein Drittel der Befragten jünger als 34 Jahre ist (29,5%, n=83), 35–49-Jährige machen 24,2% (n=68), 50–65-Jährige 35,6% (n=100) und über 65-Jährige 10,7% (n=30) der Stichprobe aus. Bezüglich des Bildungsniveaus zeigt sich ein stark akademisch geprägtes Profil: 68,0% (n=191) der Befragten verfügen über einen (Fach-)Hochschulabschluss, weitere 17,8% (n=50) haben (Fach-)Abitur. Personen mit abgeschlossener Ausbildung (8,2%, n=23) oder Haupt-/Realschulabschluss (6,0%, n=17) als höchsten Bildungsabschluss sind deutlich unterrepräsentiert. Die Einkommensverteilung weist eine breite Streuung auf: Rund 16,4% (n=41) der Befragten geben ein jährliches Haushaltsnettoeinkommen von unter 25.000 Euro an, während 14,0% (n=35) ein Einkommen über 100.000 Euro berichten. Die größte Gruppe stellt die Einkommensklasse zwischen 50.000 und 75.000 Euro mit 28,8% (n=72), gefolgt von der Klasse zwischen 25.000 und 50.000 Euro (24,0%, n=60) und 75.000–100.000 Euro (16,8%, n=42).

Die Untersuchung der Wohnsituation zeigt, dass die Verteilung der Stichprobe zwischen Eigentümern mit 52,4% (n=108) und Mietern mit 47,6% (n=98) ausgeglichen ist. Beim aktuell bewohnten Gebäudetyp dominieren Wohnungen in Mehrfamilienhäusern (MFH) mit 55,1% (n=114), gefolgt von Einfamilienhäusern (EFH) mit 21,7% (n=45), Reihenhäusern (RH) mit 12,6% (n=26) und Doppelhäusern (DHH)/Zweifamilienhäusern (ZFH) mit 10,6% (n=22). Rund 15,0% (n=30) der Befragten leben auf weniger als 50 m². Der größte Anteil mit 39,0% (n=78) hat eine Wohnfläche von 50 bis 100 m², weitere 29,5% (n=59) von 101 bis 150 m². Ein geringer Anteil von 12,0% (n=26) lebt auf 151 bis 200m² und 4,5% (n=9) der Teilnehmer haben mehr als 200 m² Wohnfläche zur Verfügung. Die Haushaltsgröße variiert ebenfalls: 21,1% (n=58) der Befragten leben allein, 41,1% (n=113) in Zwei-Personen-Haushalten und 35,3% (n=97) in Haushalten mit drei bis fünf Personen. Haushalte mit mehr als fünf Mitgliedern sind mit 2,5% (n=7) kaum vertreten.

Die aktuelle Nutzung der untersuchten Monitoringsysteme zeigt, dass die Akzeptanz beim Energiemonitoring mit 45,6% (n=128) deutlich höher ist als bei Komfortmonitoringsystemen, die aktuell von 25,0% (n=71) der Befragten genutzt werden.

3.2 Untersuchung von Nutzungsmotiven und -hemmnissen

Im ersten Teil der Untersuchung werden die Nutzungsmotive und -hemmnisse analysiert. Zu diesem Zweck wurde mithilfe einer vierstufigen Likert-Skala, die von 1 = „gar nicht“ bis 4 = „sehr stark“ reicht, der Einfluss unterschiedlicher Motive und Hemmnisse erhoben. Die abgefragten Motivationsgründe umfassten die Kontrolle des Energieverbrauchs, die Erhöhung des Wohnkomforts, die Reduktion von Energiekosten und das Einsparen von Energie. Bei den Hemmnissen wurden hohe Anschaffungs- und Wartungskosten, technische Hürden, zu geringe Energieeinsparungen, Vorbehalte gegenüber der Datensicherheit, Mangel an Informationen sowie Desinteresse an den Anwendungen abgefragt. Zur Ermittlung der Stärke des Einflusses wurde jeweils der Mittelwert verglichen. Die Interpretation der Werte erfolgt unter der Annahme, dass Mittelwerte unter 1,49 die Bedeutung „gar nicht“, Werte zwischen 1,50 und 2,49 die Bedeutung „wenig“, Werte zwischen 2,50 und 3,49 die Bedeutung „stark“ und Werte über 3,50 die Bedeutung „sehr stark“ aufweisen. Die Ergebnisse sind der Tabelle 1 zu entnehmen.

Die Ergebnisse legen nahe, dass die Kontrolle des Energieverbrauchs das Hauptmotiv für die Nutzung eines Energiemonitorings (MW=3,23; N=120; SD=,991) darstellt. Auch die Reduktion der Energiekosten (MW=3,10; N=115; SD=1,46) und das allgemeine Einsparen von Energie (MW=3,06; N=117; SD=1,028) sind starke Einflussfaktoren auf die Nutzung. Die Komforterhöhung (MW=2,40; N=114; SD=1,070) ist beim Energiemonitoring dagegen ein Faktor mit eher geringer Einflussstärke.

Tabelle 1: Mittelwertvergleich von Nutzungsmotiven und -hemmnissen (Eigene Datenerhebung, 2025)

		Energiemonitoring			Komfortmonitoring		
		MW	N	SD	MW	N	SD
Nutzungsmotive	Kontrolle des Energieverbrauchs	3,23	120	0,991	2,88	69	0,932
	Erhöhung des Wohnkomforts	2,40	114	1,070	3,38	69	0,730
	Reduktion der Energiekosten	3,10	115	1,046	2,99	69	0,833
	Einsparen von Energie	3,06	117	1,028	2,94	70	0,931
Nutzungshemmnisse	Hohe Anschaffungskosten	2,52	89	1,001	2,73	147	1,095
	Hohe Wartungskosten	1,93	84	0,916	2,20	132	1,022
	Technische Hürden	2,20	91	1,056	2,27	151	1,045
	Geringe Energieeinsparungen	2,31	88	0,998	2,47	140	1,000
	Vorbehalte ggü. Datensicherheit	2,25	95	1,091	2,31	151	1,109
	Mangel an Informationen	2,29	94	0,991	2,36	162	1,062
	Mangel an Zeit	2,04	91	0,999	1,97	157	0,930
	Kein Interesse	2,20	85	1,193	2,72	149	1,173

Beim Komfortmonitoring ist die Erhöhung des Wohnkomforts (MW=3,38, N=69; SD= ,730) dagegen das Hauptmotiv für die Nutzung. Die Reduzierung der Energiekosten (MW=2,99; N=69; SD=,833) und der Energie (MW=2,94; N=70; SD=,931) sind zwar nach wie vor starke Einflussfaktoren auf die Nutzung, werden jedoch als deutlich weniger starke Nutzungsgründe bewertet als die Erhöhung des

Wohnkomforts. Mit einem Mittelwert von 2,88 (N=69; SD=,932) stellt die Verbrauchskontrolle das geringste Motiv für die Nutzung eines komfortorientierten Monitorings dar.

Die Auswertung der Nutzungshemmnisse zeigt, dass die hohen Anschaffungskosten beim Energiemonitoring der am stärksten bewertete hemmende Faktor sind. Mit einem Mittelwert von 2,52 (N=89; SD=1,001) spiegelt dieser jedoch ein geteiltes Meinungsbild wider. Auch die anderen abgefragten Hemmnisse werden mit Mittelwerten zwischen 2,31 und 2,04 als geringfügige Hemmnisse wahrgenommen. Die hohen Wartungskosten werden mit einem Mittelwert von 1,93 (N=84; SD=,916) als das geringste Hemmnis wahrgenommen. Deutlichere Hemmnisse sind dagegen beim Komfortmonitoring erkennbar: Hier stellen die hohen Anschaffungskosten (MW=2,73; N=147; SD=1,095) und das fehlende Interesse (MW=2,72; N=149; SD=1,173) die stärksten Hemmnisse für die Nutzung dar. Zudem wird die erwartete geringe Energieeinsparung mit einem Mittelwert von 2,47 (N=140; SD=1,000) ambivalent bewertet. Die weiteren abgefragten Hürden werden mit Mittelwerten zwischen 2,36 und 2,20 als eher geringe Einflussfaktoren bewertet. Der Mangel an Zeit (MW=1,97; N=157; SD=,930) scheint aufgrund der schwächsten Bewertung das geringste Hemmnis darzustellen.

3.3 Untersuchung von soziodemografischen und wohnstrukturellen Faktoren

Im zweiten Untersuchungsschritt wurde mithilfe binär-logistischer Regressionsanalysen geprüft, welche Einflussfaktoren die Nutzung der verschiedenen Anwendungen signifikant beeinflussen. Die abhängige Variable kodiert, ob ein System im Haushalt verwendet wird (1=Nutzung, 0=keine Nutzung). In das Modell wurden alle unabhängigen Variablen als kategoriale Prädiktoren aufgenommen, darunter die soziodemografischen Faktoren Geschlecht, Altersgruppe, höchster Bildungsabschluss und Jahreshaushaltseinkommen sowie die wohnstrukturellen Faktoren Eigentumsverhältnis (Mieter/Eigentümer), Gebäudetyp, Wohnfläche und Haushaltsgröße. Die Ergebnisse sind der Tabelle 2 zu entnehmen.

Das Regressionsmodell für die Nutzung von Energiemonitoring war statistisch signifikant, χ^2(22)= 41,024; p=,008, mit einer akzeptablen Varianzaufklärung von Nagelkerkes R²=,0298, gemäß den Empfehlungen von Backhaus et al. [10]. Die Anpassungsgüte wurde mit dem Hosmer-Lemeshow-Test überprüft, der eine hohe Anpassungsgüte zeigte, χ^2(8)=7,852; p=,448. Der Gesamtprozentsatz korrekter Klassifikation war 73,6%, mit einer Sensitivität von 83,5% und einer Spezifität von 61,1%. Die Ergebnisse zeigen bei der Nutzung von Energiemonitoring signifikante Effekte für mehrere Prädiktoren. Im Vergleich zur Referenzkategorie, dem niedrigsten Bildungsabschluss (Hauptschul-/Realschulabschluss), war die Wahrscheinlichkeit einer Nutzung bei höheren Bildungsabschlüssen niedriger. Teilnehmer mit einem Fachhochschulabschluss wiesen eine signifikant niedrigere Wahrscheinlichkeit auf, ein Energiemonitoring-System zu nutzen (B=-1,838; SE=,908; p=,043). Auch der Gebäudetyp war ein signifikanter Prädiktor. Personen, die in einem Zweifamilienhaus oder einer Doppelhaushälfte leben, nutzten Energiemonitoring-Systeme signifikant weniger als Personen in Einfamilienhäusern (Referenzkategorie) (B=-1,493; SE=,709; p=,035). Die Effekte von Geschlecht, Altersgruppe, Einkommen, Eigentumsverhältnis, Wohnfläche und Haushaltsgröße waren nicht signifikant. Es zeigten sich jedoch tendenzielle Unterschiede zugunsten größerer Wohneinheiten, die in weiterführenden Analysen berücksichtigt werden könnten.

Auch das binär-logistische Regressionsmodell zur Vorhersage der Nutzung von Komfortmonitoring erwies sich als statistisch signifikant, χ^2(22)=56,040; p<,001. Die erklärte Varianz lag bei Nagelkerkes

Tabelle 2: Logistische Regression (Eigene Datenerhebung, 2025)

Nutzung	Energiemonitoring			Komfortmonitoring		
	B	SE	Sig.	B	SE	Sig.
Konstante	2,167	1,333	0,104	-2,324	1,653	0,160
Geschlecht (1 = weiblich)	-0,007	0,451	0,988	-1,194	0,709	0,092
Alter: <34 Jahre (Ref)			0,884			0,192
Alter: <35 - 49 Jahre	-0,016	0,627	0,979	-1,090	0,813	0,180
Alter: 50 - 65 Jahre	0,271	0,559	0,627	**-1,678**	**0,771**	**0,030**
Alter: >65 Jahre	0,470	0,719	0,513	-22,263	8023,436	0,998
Bildung: Haupt-/Realschulabschluss (Ref)			0,242			0,807
Bildung: Ausbildung	-1,789	1,061	0,092	-0,036	1,380	0,979
Bildung: Fachabitur	-1,856	0,978	0,058	0,759	1,156	0,511
Bildung: (Fach-)Hochschulabschluss	**-1,838**	**0,908**	**0,043**	0,697	1,107	0,529
Einkommen: <25.000 Euro (Ref)			0,198			**0,007**
Einkommen: 25.000 - 50.000 Euro	-0,402	0,621	0,518	0,139	0,914	0,879
Einkommen: 50.000 - 75.000 Euro	-0,248	0,599	0,678	1,379	0,889	0,121
Einkommen: 75.000 - 100.000 Euro	-0,080	0,706	0,910	1,926	0,996	0,053
Einkommen: >100.000 Euro	1,829	1,000	0,067	**3,553**	**1,135**	**0,002**
Eigentumsverhältnis (1 = Mieter)	-0,238	0,685	0,729	1,477	0,995	0,138
Gebäudetyp: EFH (Ref)			0,186			0,286
Gebäudetyp: ZFH/DHH	**-1,493**	**0,709**	**0,035**	0,046	0,837	0,957
Gebäudetyp: RH	-0,607	0,686	0,376	-1,306	0,943	0,166
Gebäudetyp: Wohnung in MFH	-1,040	0,703	0,139	-1,520	0,955	0,111
Wohnfläche: <50 qm (Ref)			0,523			0,091
Wohnfläche: 51 - 100 qm	0,099	0,675	0,884	0,488	0,927	0,599
Wohnfläche: 101 - 150 qm	1,022	0,881	0,246	2,197	1,179	0,063
Wohnfläche: 151 - 200 qm	0,247	1,014	0,807	1,313	1,474	0,373
Wohnfläche: >200 qm	0,774	1,312	0,555	**3,975**	**1,798**	**0,027**
Haushaltsgröße: 1 Person (Ref)			0,808			0,562
Haushaltsgröße: 2 Personen	-0,196	0,563	0,728	-0,653	0,747	0,382
Haushaltsgröße: 3 - 5 Personen	-0,578	0,651	0,375	-0,687	0,805	0,394
Haushaltsgröße: >5 Personen	-0,046	1,421	0,974	1,101	1,621	0,497
Nagelkerke R^2	-	-	0,298	-	-	0,438

R^2=,438, die als gut zu bewerten ist. Die Modellanpassung wurde mithilfe des Hosmer-Lemeshow-Tests geprüft, $\chi^2(8)$ = 5,904; p=,658. Die Klassifikationsgenauigkeit des Modells betrug insgesamt 82,9%, wobei die Sensitivität bei 92,9% und die Spezifität bei 50,0% lag.

Signifikante Effekte zeigten sich bei der Nutzung von Komfortmonitoring für die Alters- und Einkommenskategorien. Personen im Alter von 50 bis 65 Jahre nutzen Komfortmonitoring-Systeme seltener als Personen die jünger als 34 Jahre sind (Referenzkategorie) (B=-1,678; SE=,771; p=,030). Auch in den anderen Altersgruppen ist dieser Trend zu erkennen. Er konnte jedoch nicht signifikant nachgewiesen werden. Ebenso war ein höheres Haushaltseinkommen (>150.000 Euro) im Vergleich zur Referenz, der niedrigsten Einkommensklasse (<25.000 Euro) signifikant positiv assoziiert mit der Nutzung (B=3,553; SE=1,135; p=,002). Zudem war die Wohnfläche ein signifikanter Prädiktor: Haushalte mit mehr als 200 m² Wohnfläche wiesen eine signifikant höhere Wahrscheinlichkeit der Nutzung auf als

Haushalte mit weniger als 50 m² (Referenzkategorie), B=3,975; SE=1,798; p=,027. Kein signifikanter Einfluss zeigte sich hingegen für Geschlecht, Bildungsabschluss, Eigentumsverhältnis, Gebäudetyp oder Haushaltsgröße.

4 Zusammenfassung und Diskussion

Die Untersuchung zeigt, dass die Akzeptanz und Nutzung von Energiemonitoring in der Stichprobe deutlich höher ausfällt als die von Komfortmonitoring. Zentrale Nutzungsmotive für energiebezogene Systeme sind die Kontrolle des Verbrauchs, die Steigerung der Energieeffizienz und die Reduzierung der Energiekosten. Komfortmonitoring wird hingegen primär durch den Wunsch nach mehr Wohnkomfort motiviert. Die Ergebnisse legen nahe, dass Energieeffizienz und Komfortsteigerung nicht im Widerspruch stehen, sondern je nach Anwendung unterschiedlich gewichtet werden. Hemmnisse wie technische Hürden, Datenschutzbedenken oder Informationsdefizite wurden von den Befragten mehrheitlich als gering eingeschätzt. Die Anschaffungskosten stellten jedoch für beide Systeme eine Barriere dar. Auffällig ist zudem das relativ hohe Desinteresse gegenüber dem Komfortmonitoring. Dies könnte auf eine geringe wahrgenommene Alltagsrelevanz oder mangelnde Aufklärung hinweisen.

Die binär-logistischen Regressionsanalysen identifizieren signifikante Zusammenhänge zwischen der Nutzung und soziodemografischen bzw. wohnstrukturellen Merkmalen. So wird Energiemonitoring beispielsweise häufiger in Einfamilienhäusern und von Personen mit niedrigeren Bildungsabschlüssen genutzt. Die Nutzung von Komfortmonitoring korreliert hingegen positiv mit Einkommen, Wohnfläche und jüngerem Alter. Dies lässt auf eine stärkere Verknüpfung mit Lebensstilfaktoren und Technikaffinität schließen. Diese Ergebnisse unterstreichen die Relevanz zielgruppenspezifischer Förder- und Kommunikationsstrategien. Während Energiemonitoring bereits über ein breiteres funktionales Akzeptanzpotential verfügt, bedarf es für Komfortmonitoring einer differenzierten Kommunikations- und Aufklärungsstrategie, die die subjektive Relevanz und den Mehrwert stärker adressiert. Politik und Praxis sind somit aufgefordert, neben der technologischen Entwicklung auch die soziale Anschlussfähigkeit digitaler Monitoringsysteme systematisch zu stärken.

Die Generalisierbarkeit der Ergebnisse ist jedoch aufgrund einer selektiven Stichprobenzusammensetzung (u. a. hoher Männeranteil, bildungs- und einkommensstarke Haushalte) eingeschränkt. Die überrepräsentierte Teilnehmerstruktur könnte Verzerrungen bei den Effekten von Bildung und Einkommen verursachen und so die Abweichungen von bisherigen Forschungsergebnissen erklären. Künftige Studien sollten diese gegenläufigen Zusammenhänge mit repräsentativeren Stichproben vertiefen. Dennoch leistet die Studie einen wichtigen Beitrag zum Verständnis zielgruppenspezifischer Dynamiken im Kontext der digitalen Transformation im privaten Wohnumfeld.

Danksagung

Diese Arbeit entstand im vom Bundesministerium für Wirtschaft und Energie (BMWE) im Rahmen des 7. Energieforschungsprogramms geförderten Reallabors der Energiewende *SmartQuart* (FKZ: 03EWR010B). Wir danken allen Projektpartnern sowie den teilnehmenden Haushalten für ihr Mitwirken an der Datenerhebung. Außerdem bedanken wir uns bei den anonymen Gutachtern für ihre hilfreichen Anmerkungen.

Literatur

[1] Bundesministerium für Wirtschaft und Klimaschutz, *Energieeffizienz in Zahlen*, 2023. Adresse: https://www.bmwk.de/Redaktion/DE/Publikationen/Energie/energieeffizienz-in-zahlen-2023.html

[2] D. Brounen, N. Kok und J. M. Quigley, »Energy literacy, awareness, and conservation behavior of residential households«, *Energy Economics*, Jg. 38, S. 42–50, 2013. DOI: https://doi.org/10.1016/j.eneco.2013.02.008 Adresse: https://www.sciencedirect.com/science/article/pii/S0140988313000327

[3] H. Allcott und T. Rogers, »The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation«, *American Economic Review*, Jg. 104, Nr. 10, S. 3003–37, Okt. 2014. DOI: 10.1257/aer.104.10.3003 Adresse: https://www.aeaweb.org/articles?id=10.1257/aer.104.10.3003

[4] E. Venturi, F. Ochs und G. Dermentzis, »Identifying the influence of user behaviour on building energy consumption based on model-based analysis of in-situ monitoring data«, *Journal of Building Engineering*, Jg. 64, 2023. DOI: 10.1016/j.jobe.2022.105717

[5] J. Froehlich, L. Findlater und J. Landay, »The design of eco-feedback technology«, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Ser. CHI '10, Atlanta, Georgia, USA: Association for Computing Machinery, 2010, S. 1999–2008. DOI: 10.1145/1753326.1753629 Adresse: https://doi.org/10.1145/1753326.1753629

[6] Ó. García, J. Prieto, R. Alonso und J. Corchado, »A Framework to Improve Energy Efficient Behaviour at Home through Activity and Context Monitoring«, *Sensors*, Jg. 17, Nr. 8, 2017. DOI: 10.3390/s17081749

[7] O. Iweka, S. Liu, A. Shukla und D. Yan, »Energy and behaviour at home: A review of intervention methods and practices«, *Energy Research & Social Science*, Jg. 57, 2019. DOI: 10.1016/j.erss.2019.101238

[8] Wissenschaftliche Begleitforschung Energiewendebauen - Modul 2, »Nutzendenverhalten – Einflussfaktoren und Einflussnahme auf die Gebäudeenergieperformance. Querauswertung der Projekte der Förderinitiative.«, Techn. Ber., 2024. Adresse: https://www.energieforschung.de/lw_resource/datapool/systemfiles/agent/ewbpublications/f9f8a7fe-7fd7-11ef-aa67-a0369fe1b534/live/document/Auswertung_Nutzendenverhalten.pdf

[9] H. Shamon u. a., »Smart Energy in Haushalten: Technologien, Geschäftsmodelle, Akzeptanz und Wirtschaftlichkeit«, Forschungszentrum Jülich, Report, 2021. Adresse: https://epub.wupperinst.org/frontdoor/deliver/index/docId/7822/file/7822_Smart_Energy.pdf

[10] K. Backhaus, B. Erichson, S. Gensler, R. Weiber und T. Weiber, *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*. Berlin: Springer Gabler, 2023, Bd. 17. Aufl.

The Role of AI Agents in Construction Project Management

Anja Brelih[1], Marko Robnik Šikonja[2] and Robert Klinc[1]
[1]University of Ljubljana, Faculty of Civil and Geodetic Engineering, Slovenia
[2]University of Ljubljana, Faculty of Computer and Information Science, Slovenia
E-mail(s): anja.brelih@fgg.uni-lj.si, marko.robniksikonja@fri.uni-lj.si, robert.klinc@fgg.uni-lj.si

Abstract: The integration of Large Language Models (LLMs) into Artificial Intelligence (AI) agents presents new opportunities for enhancing construction project management through improved collaboration, reasoning, and workflow automation. This paper investigates how LLM-powered AI agents can support complex construction tasks such as scheduling, resource allocation, risk assessment, and information coordination. We conduct a structured review of recent literature across construction informatics, AI-driven project management, and agent-based systems to identify trends, capabilities, and limitations. Our findings show that while AI agents offer strong potential to assist project teams with dynamic decision-making and routine task automation, their effectiveness is limited by data quality, domain adaptation, and the stochastic nature of LLMs. These insights are significant for researchers and practitioners aiming to introduce AI agents into real-world construction workflows, where reliability, explainability, and human oversight are essential. The paper highlights the need for further exploration of Human-in-the-Loop (HITL) designs and calls for standardisation of agent capabilities to ensure safe, transparent, and practical deployment in the construction industry.

Keywords: AI Agents, LLM, Project Management, Construction Workflows, Construction Informatics

DOI: 10.18154/RWTH-CONV-254903. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Construction project management is increasingly shaped by digital transformation, driven by the need for greater adaptability and efficiency. Among the technologies reshaping this field, artificial intelligence (AI) has gained momentum for its ability to support decision-making, automate repetitive tasks and improve project outcomes in planning and monitoring. Construction projects are characterised by unique products that require a unique sequence of processes and group of partners [1], making standardised automation difficult. In this context, the emergence of AI agents, autonomous software units that are able to perceive their environment, reason, and execute tasks, offers new opportunities to manage the complexity and variability of construction workflows [2]. Unlike conventional systems, AI agents can learn from data, adapt to new constraints and collaborate with human stakeholders. Their potential is significantly enhanced by the integration of Large Language Models (LLMs), which enable advanced understanding of natural language, multi-step reasoning and interaction with external tools.

LLM-powered AI agents can act as intelligent collaborators that process diverse information, identify project tasks and facilitate communication [3], [4].

This paper explores the role of AI agents in construction project management. The rest of the paper is organised as follows. Section 2 outlines the theoretical and technological foundations, including the development of AI agents and construction informatics. Section 3 discusses the applicability of AI agents in construction management, covering their capabilities, suitability for specific tasks and integration into existing workflows. Section 4 addresses key challenges and limitations such as data security, trust and adaptability. Finally, Section 5 concludes the paper and identifies directions for future research and practical implementation.

2 Background and related work

This section outlines the theoretical foundations of AI agents in construction, grounded in construction informatics. It also traces the evolution from Construction 3.0 to 5.0, highlighting the growing role of LLMs and current research gaps.

2.1 Construction informatics and project management

Construction informatics addresses the unique characteristics of construction data, such as its fragmentation, its temporality and its project-specific context. As Turk [1] noted, the field investigates the processes of creating, processing and communicating construction-specific information, whether in human or machine-readable form. These processes are traditionally based on structured input-output relationships, with “glue” processes connecting information generation and use through static data formats, defined APIs or hard-coded scripts. Figure 1 illustrates this workflow structure and highlights the rigid, static nature of traditional data exchanges.

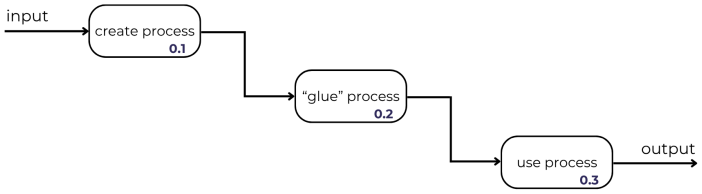


Figure 1: Construction Informatics Workflow [1].

Construction project management, while traditionally focused on time, cost and quality, is increasingly expanding to include environmental and social sustainability objectives [5]. In this broader context, digital tools, especially AI, are becoming increasingly important as decision support systems. AI agents, for example, are developing into valuable collaborators and can contribute to:

- **Automated monitoring and reporting:** Agents continuously assess alignment with schedule and budget constraints and can suggest corrective measures in the event of deviations [6].
- **Decision support:** By synthesising diverse data sources, agents can create alternative scenarios and help choose optimal project paths [7].
- **Risk forecasting:** By analysing current project statuses and historical data in real time, agents can identify emerging risks and suggest risk mitigation strategies [5].

These AI applications are well suited to the increasing complexity and unpredictability of modern construction projects. As such, construction informatics provides not only the technological foundation, but also the methodological framing for integrating AI agents into real construction workflows. While automated workflows follow fixed predefined steps, AI agents dynamically plan, execute actions with tools, and reflect on results (Figure 2).

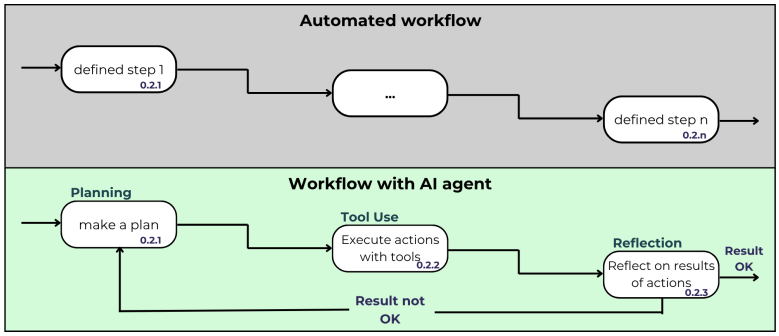


Figure 2: Comparison between Automated workflow and Workflow with AI agent. Adapted from [8].

2.2 From Construction 3.0 to 5.0

The construction industry has started to integrate LLMs and AI agents into project management workflows (Figure 2), which is an important step towards the Construction 5.0 paradigm. While Construction 3.0 is mainly about the separate use of digital tools, where a human interacts with a digital environment separate from the physical environment via a computer, Construction 4.0 is the digital transformation of the construction industry, i.e. the integration of the physical and digital environment through cyber-physical systems [9]. This creates a connected ecosystem between machines, data and the user. As an extension of the Industry 5.0 concept, Construction 5.0 focuses on people again and the development of intelligent systems that work together with people instead of replacing them, or in other words, from machine-to-machine integration back to human-to-machine integration [10]. In such environment, the AI agent does not take full control, but acts as a collaborator – with a clearly limited role, supervised by a human. This approach follows the Human-in-the-Loop (HITL) principle, which is one of the foundations of Industry 5.0. The connected ecosystem is further deepened in Construction 5.0 through the introduction of cognitive cyber-physical systems, where technologies not only monitor and control processes, but also understand context, collaborate with people and adapt to their needs.

2.3 AI Agents and Large Language Models

Research on AI agents is increasingly focusing on their integration into collaborative environments where they act as virtual team members. In this paper, we define AI agents as systems capable of autonomous task execution, contextual decision-making, and interaction with tools or users. Studies suggest that well-trained and accepted agents can improve information coordination and reduce conflicts in project teams. However, mistakes made by AI agents are often viewed more critically by users than those made by human colleagues [11].

With the rapid development of LLMs, a new generation of AI agents has emerged. These agents are able to understand natural language and generate complex, context-appropriate responses [3].

LLMs enable agents to interpret user input, perform structured reasoning using techniques such as Chain-of-Thought or Tree-of-Thoughts and interact with external tools and data sources via APIs [4], [6]. Additional relevant capabilities include processing images with Visual Language Models (VLMs) and structured inputs such as JSON and XML.

Recent developments include Large Reasoning Models (LRMs), which focus on strategic planning and high-level decision-making. In combination, LLMs and LRMs allow agents to execute tasks while also reflecting on their outcomes, significantly expanding the range of complex problems they can solve [12], [13]. In construction management, such agents are increasingly relevant for decision support, task assignment, risk prediction and communication between stakeholders [7], [11].

2.4 Research gap

Despite rapid advancements in the development and application of AI agents, several challenges remain unaddressed in the context of construction project management. While AI agents offer significant advantages in managing complex and dynamic workflows, they are not universally suitable for all project contexts. This highlights the need to first assess when and where the use of AI agents is feasible. Second, there is a lack of standardized approaches for integrating AI agents into construction-specific workflows, which often differ from other industries due to their one-of-a-kind nature. Third, while AI agents show potential for augmenting teamwork, research on human–AI collaboration (particularly in high-stakes, safety-critical environments like construction) is still limited. Additionally, issues of security, privacy, and explainability are insufficiently explored, especially with regard to sensitive project data and the traceability of agent-generated decisions [2]. Finally, although Construction 5.0 emphasizes values such as sustainability, human-centricity, and trust, these principles are not yet fully embedded in the design and deployment of AI agents [7]. These gaps underscore the need for a conceptual and methodological framework that supports the development of AI agents for construction project management.

3 AI Agents in Construction Management

AI agents enable natural language interaction, intelligent task automation and predictive real-time analyses. LLM-based agents enable construction professionals to engage in dialogue with systems and support both planning and tool usage through intuitive interfaces [4], [6]. Advances in reasoning capabilities, processing of images and structured data and formal language integration improve the ability of agents to accurately interpret complex construction tasks [12], [13]. In task automation, AI agents streamline project planning, resource allocation and risk assessment by dynamically adapting teams to project requirements and identifying potential bottlenecks at an early stage [5], [7]. Through workflows with AI agents, they adapt to the changing conditions on the construction site and thus meet the goals of the digital transformation of Construction 4.0 and 5.0 [8], [9], [10].

3.1 AI agents suitable tasks

Before using AI agents in construction workflows, it's essential to ask a fundamental question: *Is this task suitable for an AI agent, or would another solution be better?* Agents should not be seen as a one-size-fits-all solution. They are most valuable in workflows that are complex, dynamic and require interaction with multiple tools. However, their use must be carefully considered based on task characteristics and project risk.

As shown in Figure 3, there are several decision points that determine whether an AI agent is appropriate. This decision flow evaluates the task across multiple dimensions and helps to decide whether an AI agent, a simpler LLM-based solution or classic automation is more appropriate.

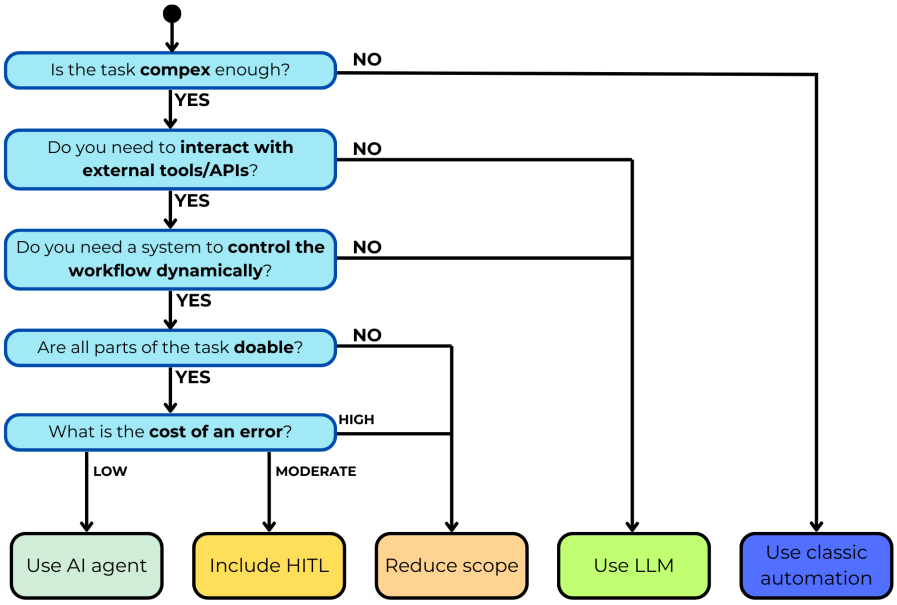


Figure 3: Task suitability evaluation for AI agents.

- **Task complexity**: If the task is not complex enough to require autonomous reasoning, a rule-based method or classic automation may suffice. Conversely, tasks that involve uncertain outcomes, coordination or adaptive planning may benefit from agent-based solutions.
- **Text-only and tool interaction**: If the task only involves creating or summarising text, a standard LLM is sufficient. However, if the task requires interaction with external APIs or tools, an agent may be needed.
- **Workflow control**: AI agents should only be considered if the system needs to dynamically control the workflow based on real-time input. Static or linear tasks do not justify the complexity of agents and their potential unreliability and stohastic behaviour.
- **Feasibility**: If not all parts of the task are yet automatable, the scope should be reduced or partially delegated. Tasks that remain too ambiguous should not be assigned to agents. In this case, the scope should be reduced or limited to what the agent can handle.
- **Error tolerance**: In safety-critical scenarios or scenarios where the stakes are high, as is typical in construction, humans must remain in control. Tasks with low error costs can be fully automated, but for tasks with moderate risk, agents should only support humans and not decide autonomously.

A full AI agent integrates memory, tool utilisation, context retrieval and workflow control. This makes them powerful, but also risky if used incorrectly. Figure 3 illustrates that many tasks in construction do not justify the use of a full agent and can be better handled with LLMs. In most construction scenarios, simpler systems with HITL are more suitable than fully autonomous agents, as they need to be accountable and explainable. Therefore, AI agents in construction should act as contributors

rather than decision makers. They are most effective when they offer alternative solutions and support, while leaving the final decision to engineers. HITL should not only be a practical safeguard, but a necessary design principle in the field of construction (e.g. audit logs of agents, checkpoints for human overrides, thresholds for confidence in decisions). AI agents should be designed from the outset to enable traceability and validation by engineers at key decision points.

3.2 Integration with Construction Workflows

AI agents can be integrated into construction workflows to enhance communication, decision-making, and task execution. This section briefly presents interaction scenarios with stakeholders and illustrates possible use cases in construction project environments.

3.2.1 Communication with stakeholders

AI agents equipped with LLMs and VLMs enable intuitive and context-aware communication between systems and stakeholders. Instead of requiring structured input formats, users can interact with agents via queries (e.g. ask for project status, request scenario comparisons, delegate task updates). This lowers the barrier to information access and helps bridge fragmented communication across roles and disciplines. For example, a site manager could ask an agent for real-time material delivery forecasts or potential delays and receive a response based on integrated project data.

In addition, agents can translate technical information into appropriate formats for stakeholders. When communicating with non-technical personnel or clients, agents can create customised summaries that reflect the relevant information. This supports transparency and helps align diverse expectations on construction projects.

3.2.2 Example workflows or theoretical use cases

In practise, AI agents can be integrated into several important construction workflows:

- **Schedule coordination**: Agents can automatically update project schedules based on current progress reports or external disruptions, and can also simulate alternative schedules when constraints change. For instance, in a disrupted schedule coordination scenario, an AI agent could propose a sequence swap based on material delays and flag it to a human supervisor for validation.
- **Resource allocation**: AI agents can suggest resource shifts based on evolving site needs, helping teams respond more quickly to shortages or bottlenecks.
- **Risk assessment**: Agents can continuously scan inputs such as progress logs, cost reports or environmental conditions to identify early warning signs of potential risk.
- **Documentation and reporting**: Agents can automatically generate site reports, meeting summaries or compliance logs, improving administrative efficiency and reducing human error.

These examples show how workflows with AI agents differ from traditional digital tools. Instead of executing pre-programmed steps, agents monitor progress, select contextual actions and dynamically adapt the workflow [8]. This makes them particularly valuable in the fragmented, evolving conditions typical of construction.

4 Challenges and limitations

Despite the promising capabilities of AI agents, their use in construction project management raises some challenges that need to be addressed to ensure responsible and effective adoption.

4.1 Data security and privacy

Construction projects contain sensitive data including contracts, costs, personnel records and safety documentation. When AI agents process such data (often via APIs or cloud services), there is a risk of data leakage or misuse. Deng et al. [2] emphasise that AI agents are particularly vulnerable to the unpredictability of multi-step user inputs, the complexity of internal executions, the variability of operational environments and interaction with untrusted external entities. This risk is magnified in construction, a highly regulated industry. Clear architectural guidelines for secure data flows and on-premise deployment options should be considered to meet compliance requirements.

4.2 Bias, transparency, and trust

AI agents that rely on LLMs inherit the limitations of these models, including possible biases in the training data which can influence decision-making. One of the most pressing challenges is the explainability of agent decisions. In domains where the stakes are high, users need to understand not only the outcomes, but also the reasoning behind them. As Dennis et al. [11] point out, human trust in AI agents relies heavily on perceived clarity. Designing agents with built-in mechanisms for explanation, oversight and human override is critical.

4.3 Adaptability to dynamic project environments

Construction projects are dynamic by nature and often require rapid re-planning due to design changes, weather conditions or resource availability. While AI agents show promise in managing this complexity through contextual reasoning and workflow adaptation, their effectiveness in practise is limited by generalisation issues and an insufficient foundation of project-specific data. Agents trained on generic corpora may struggle to adapt to the unique terminology and evolving requirements of individual sites. Their stochastic nature raises concerns about reliability, reproducibility, and deployment risks, especially for consistent behavior in real-world applications. Without customised integration and domain adaptation (as highlighted by Taboada et al. [5]), AI agents run the risk of making irrelevant or impractical suggestions. A HITL approach, where agents support decision making but engineers remain in control, offers a more resilient solution that meets the needs of construction domain.

5 Final remarks

AI agents offer a promising new dimension to construction project management by improving adaptability, communication and decision support. This paper highlights their potential role, their current limitations and the importance of careful task selection. To move forward, standardisation, HITL integration and domain-specific adaptation are essential. As we move towards Construction 5.0, future research should focus on practical implementation strategies that align AI agents with human supervision and develop common benchmarks for evaluating the trustworthiness and effectiveness of agents in construction projects.

Acknowledgements

This research was supported by the Slovenian Research and Innovation Agency (ARIS) under the Young Researcher funding program and research programs E-Construction (P2-0210) and Language Resources and Technologies for Slovene (P6-0411).

References

[1] Ž. Turk, "Construction informatics: Definition and ontology", *Advanced engineering informatics*, vol. 20, pp. 187–199, 2006.

[2] Z. Deng et al., "AI Agents Under Threat: A Survey of Key Security Challenges and Future Pathways", *ACM Computing Surveys*, vol. 57, no. 7, 2025. DOI: 10.1145/3716628

[3] Z. Xi et al., "The rise and potential of large language model based agents: A survey", en, *Science China Information Sciences*, vol. 68, no. 2, 2025.

[4] Z. Liu et al., *AgentLite: A Lightweight Library for Building and Advancing Task-Oriented LLM Agent System*, 2024. arXiv: 2402.15538.

[5] I. Taboada, A. Daneshpajouh, N. Toledo, and T. de Vass, "Artificial Intelligence Enabled Project Management: A Systematic Literature Review", *Applied Sciences*, vol. 13, no. 8, 2023. DOI: 10.3390/app13085014

[6] J. Ruan et al., "TPTU: Large Language Model-based AI Agents for Task Planning and Tool Usage", in *NeurIPS 2023 Foundation Models for Decision Making Workshop*, 2023.

[7] N. Mikhridinova, F. A. Yurdakul, and C. Wolff, "Using Large Language Models for Project Staffing: Evaluation of GPT-Based Mapping of Teams to Projects", in *Proceedings of the 12th IPMA Research Conference*, 2024, pp. 1–19. DOI: 10.56889/cykr4175

[8] M. Newhauser, L. Monigatti, P. Yadav, and T. Çelik, *What Are Agentic Workflows?*, Mar. 2025. [Online]. Available: https://weaviate.io/blog/what-are-agentic-workflows

[9] R. Klinc and Ž. Turk, "Construction 4.0 - digital transformation of one of the oldest industries", *Economic and business review*, vol. 21, no. 3, pp. 393–410, 2019.

[10] M. A. Musarat, M. Irfan, W. S. Alaloul, A. Maqsoom, and M. Ghufran, "A Review on the Way Forward in Construction through Industrial Revolution 5.0", *Sustainability*, vol. 15, no. 18, 2023. DOI: 10.3390/su151813862

[11] A. R. Dennis, A. Lakhiwal, and A. Sachdeva, "AI Agents as Team Members: Effects on Satisfaction, Conflict, Trustworthiness, and Willingness to Work With", *Journal of Management Information Systems*, vol. 40, no. 2, pp. 307–337, 2023. DOI: 10.1080/07421222.2023.2196773

[12] Z. Li, W. Hua, H. Wang, H. Zhu, and Y. Zhang, *Formal-LLM: Integrating Formal Language and Natural Language for Controllable LLM-based Agents*, 2024. arXiv: 2402.00798.

[13] X. Zhou et al., *Exploring the Necessity of Reasoning in LLM-based Agent Scenarios*, 2025. arXiv: 2503.11074.

Automated Extraction of Information Requirements from Building Regulations Using LLMs and Ontologies

Sven Zentgraf[1] and Leonie Zimmermann[1]
[1]Chair of Computing in Engineering, Ruhr-Universität Bochum, Bochum, Germany
E-mail(s): sven.zentgraf@rub.de, leonie.zimmermann@rub.de

Abstract: In the context of building requirements checking, the frequent process of information requirement extraction, even in digital Building Information Modeling (BIM) processes, involves a high amount of manual extraction of information requirements from non-machine-readable regulatory documents. This process is both time-consuming and prone to inconsistencies and errors. This poses significant challenges to efficient data processing and validation. Therefore, a method is needed that simplifies the process of extracting information requirements and deploying this information in a machine-readable format. To address these limitations, this paper proposes an automated approach that leverages a Large Language Model (LLM) and prompts engineering to extract relevant information requirements from building guidelines. The extracted information is then structured according to the pre-defined ISOProps Ontology as a corresponding A-Box ontology. The syntactic and semantic correctness of the generated A-box ontology is validated with the help of Shapes Constraint Language (SHACL) shapes generated from the associated T-box ontology. This ensures formal correctness and conformity with the ISOProps ontology. A human-in-the-loop approach qualitatively verifies the technical correctness. The feasibility of this approach was demonstrated on sample documents, highlighting its potential for streamlining the digital representation of building requirements.

Keywords: Building information modeling (BIM), Large Language Model (LLM), Prompt Engineering, Information Requirement Extraction, A-Box ontology

DOI: 10.18154/RWTH-CONV-254904. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

In the field of building requirements checking, the extraction of information requirements from regulatory documents remains a significant challenge, even within digital BIM processes. Typically, this task involves a considerable amount of manual effort, as the relevant information is often embedded in non-machine-readable documents. This manual extraction is not only time-consuming but also susceptible to inconsistencies and errors, which can hinder efficient data processing and validation [1].

To overcome these challenges, there is a clear need for a method that both simplifies the extraction of information requirements and enables their deployment in a machine-readable format [2], [3]. This paper introduces an automated approach designed to address these limitations by leveraging an LLM and prompt engineering techniques to extract relevant information from building guidelines. This involves not only identifying explicit requirements but also interpreting the implicit knowledge and contextual dependencies embedded in the text. The extracted requirements are then structured according to the ISOProps Ontology, as an A-Box ontology. The ISOProps Ontology is built upon the data model of ISO 23386, which provides standardized methods for describing, creating, and maintaining building-related information requirements as properties and property groups [4], [5].

To ensure both syntactic and semantic correctness, the generated A-Box ontology is validated using SHACL shapes derived from the associated T-Box ontology, guaranteeing formal compliance with the ISOProps Ontology. Additionally, the technical accuracy of the extracted information is qualitatively assessed through a human-in-the-loop approach. The feasibility and effectiveness of this automated method are demonstrated on sample regulatory documents, underscoring its potential to streamline the digital representation and validation of building requirements.

2 Background

2.1 Information Requirements checking for Regulatory Compliance

Regulatory compliance checking in the Architecture, Engineering, and Construction (AEC) industry is the process of verifying that a building design or BIM model adheres to applicable codes, standards, and regulations. Traditionally, this process has been labor-intensive as it is performed predominantly by human experts, who interpret written regulations and manually inspect drawings or BIM models for violations. In recent years, there has been a strong drive to automate regulatory checking using BIM, in order to reduce errors, save time, and ensure that complex regulations are consistently applied [1, 6]. Early implementations of Automated Code Compliance showed promise in improving design quality and shortening permitting times [7]. However, automating compliance checking is challenging because building regulations are often lengthy, complex, and sometimes ambiguous [8]. Over the last decade, multiple approaches have been researched and developed to represent regulatory information requirements and rules in a digital format and to implement rule-checking engines [6], [7], [8]. At a high level, these approaches can be grouped into a few categories [8]. These categories range from rule-based, proprietary BIM model checkers and standard-based schemas to model view definitions, as well as approaches based on the Semantic Web and linked data, or natural language processing and structured text. The approach employed in this study is based on Semantic Web and Linked Data technologies, relying on the IR-Ontology network [4], which is further introduced in Section 2.2. Pauwels et al. [9] presents a related work introducing a semantic rule-checking system where various technical requirements from regulations are encoded as SHACL rules and OWL axioms. Their results indicate that a large portion of regulatory requirements can be validated automatically using Semantic Web reasoning [9]. Another example is the work by Nuyts et al. [8] who used OWL and SWRL to encode parts of the Flemish building regulations and checked models with a reasoner. In addition, a comparative study was conducted that utilized the same set of regulatory requirements in eight

different ways, employing two proprietary/IFC-based tools, two generic data schema approaches, and four Semantic Web approaches.

2.2 Digital Standards and Ontologies for Information Requirements

The RE-ING are central technical guidelines issued by the German Federal Ministry for Digital and Transport (BMDV) that regulate the planning and design of engineering structures in federal infrastructure projects. In practice, they provide detailed specifications that go beyond the scope of general standards such as the Eurocodes and DIN norms, particularly concerning application-specific requirements for structures like bridges, retaining walls, and noise barriers [10]. As part of ongoing digitalization efforts, there are initiatives to convert RE-ING into machine-interpretable formats, such as XML compliant with the NISO Standard Tag Suite (NISO STS), a standardized XML format for encoding and exchanging technical and scientific documents. This is aimed at enabling automated processing and integration with BIM systems.

In parallel to these transformations, broader standardization initiatives are emerging that focus on formalizing information requirements. Notably, various methodologies have been developed to structure the exchange and validation of information throughout project lifecycles. Among these, the Information Delivery Manual (IDM)[1] and the Information Delivery Specification (IDS)[2] are promoted by buildingSMART International. However, despite the growing theoretical foundation, practical implementation often lags. Many projects lack clearly defined, machine-readable information requirements, which hinders the ability to automate the validation of delivered data.

To address these challenges, recent research has examined how semantic technologies such as Linked Data and ontologies can support a more rigorous and interoperable management of information requirements. A notable contribution to the digital management of information requirements in construction is presented by Filardo et al. [4]. An aligned ontology network that harmonizes the LOIN (Level of Information Need), Data Templates (DT), and ISOProps ontologies was introduced. It demonstrates how these ontologies can be coordinated to consistently define properties and construction objects, following international standards such as ISO 23386, ISO 23387, and ISO 7817. This modular and reusable ontology network supports the interoperability and scalability of digital workflows in construction, offering a flexible foundation for the automated validation and provision of project-specific information [4].

2.3 Large Language Models for Information Extraction

LLMs are advanced neural networks trained on extensive textual data to learn language patterns. After a pre-training phase, they can perform various language-based tasks such as translation, answering questions, and summarizing content [11].

In recent years, LLMs have rapidly gained attention, driven by significant advances in model architectures and training on bigger datasets comprising billions of parameters. Their performance has improved, enabling them to handle a broad spectrum of tasks, mainly depending on the diversity and quality of the data they are exposed to during training. State-of-the-art models, such as GPT-4o,

[1] https://technical.buildingsmart.org/standards/information-delivery-manual/
[2] https://technical.buildingsmart.org/projects/information-delivery-specification-ids/

exemplify this progress by extending beyond text generation to include capabilities like code synthesis, image generation, and the processing of audio and video inputs [12].

The Mistral LLM, developed by Mistral AI, is an open-source model that offers state-of-the-art language understanding and generation. Its public release provides researchers and developers with free access to a powerful tool for advanced NLP applications [13].

As these models become increasingly capable and accessible, the challenge shifts to effectively directing these models. This is where prompt engineering plays a central role. Prompt engineering is the practice of designing effective inputs to guide LLMs toward producing accurate, relevant, and domain-specific outputs. It has emerged as a crucial discipline for harnessing the full capabilities of LLMs across diverse tasks. Two foundational paradigms in this area are zero-shot and few-shot prompting. Zero-shot prompting in this regard refers to the process where an LLM is given a task description or instruction in natural language. Still, no explicit examples of the task are provided within the prompt. The model relies entirely on its pre-trained knowledge and generalization capabilities to interpret and perform the requested task. This approach is particularly advantageous in scenarios where labeled data is scarce or unavailable, as it enables the immediate application of LLMs to new domains without requiring additional training or annotation efforts. Few-shot prompting, in contrast, involves including a small number of input-output examples within the prompt, alongside the task instruction.

This technique is beneficial when some representative examples are available, but not enough to justify full-scale supervised fine-tuning [14].

Höltgen et al. [15] explore the capabilities of LLMs for semantic data integration, focusing on evaluating LLMs for generating R2RML (RDB to RDF Mapping Language) mappings [15]. Recent research also demonstrates the practical value of LLMs for information extraction in the context of automated compliance checking. For example, Iversen [16] from NTNU developed and evaluated an LLM-based artifact for validating building regulations against BIM models. The study focused on extracting and classifying regulatory requirements from natural language documents, identifying dependencies between rules, and mapping these requirements to BIM data for compliance assessment [16].

3 Methods

This study examines the application of LLMs for the automated conversion of regulatory documents into an ontology-based representation. Specifically, prompt engineering techniques are employed to guide an LLM in converting RE-ING text passages into A-box statements, grounded in the T-box defined by the ISOProps ontology network proposed by Filardo et al. [4].

The transformation process is implemented via a Python-based toolchain that interacts with the Mistral API. Within the prompts, a series of curated examples is included that demonstrate how textual guideline content should be interpreted and formalized as individual ontology instances. By iteratively adapting the prompt and refining examples, the accuracy and consistency of the generated A-box ontology are improved. To support this conversion task, we adopt a few-shot prompt engineering strategy. The prompt is carefully structured to guide the model in generating an ontology from

construction-related requirements by breaking the task into distinct, manageable components (cf. Table 1).

Table 1: Structured prompt for generating an ontology from construction-related texts.

Category	Description
Objective	Generate an ontology from construction-related requirements
Input	XML document: `{text}`
Task	- Extract explicit construction requirements - Map terms (only those found in the text) - Generate RDF triples
Type Definitions	- Group of Properties: `{group_of_properties_definition}` - Property: `{property_definition}`
Triple Generation	- Create for each property and property group - Format: Turtle (RDF) - Groups include category and definition - Properties reference their group
Additional Rules	- Do not invent new terms - Use `DefiningValuesList` for dependencies - Mark AI-generated definitions clearly
Example File	`{example_file}`
Expected Output	Turtle file with all generated RDF triples only

The prompt introduces the LLM as an expert in the analysis of construction-related documents. It provides clear instructions on how to extract entities of type *Group of Properties* and *Property*, using definitions aligned with the ISOProps ontology. It starts with a clear objective and input specification, ensuring the model understands the domain and expected data format. The task is divided into three steps: requirement extraction, term mapping, and generating RDF triples. Type definitions and output constraints are explicitly included to enforce semantic consistency and ensure compliance with formal Terse RDF Triple Language (Turtle) syntax. Additional rules prevent the model from inventing terms and ensure transparency in generated content. A detailed Turtle serialization schema is included, specifying how output triples should be structured, along with an annotated example derived from a real RE-ING section. This enables the LLM to learn both the structural and semantic patterns required for the ontology population.

```
tempo-usecase:propertyGroupNameAlsEinSubstantiv
       rdf:type isoprops:GroupOfProperties ;
       isoprops:categoryOfGroupOfProperties "Kategorisierung hier einfügen" ;
       tempo:definitionInLanguage "(KI-generiert) Wenn keine definition im Text steht, generiere eine
           und setze (KI-generiert) davor. Ansonsten entnehme die dem Text."@de ;
       tempo:nameInLanguage "Name der Gruppe von Eingeschaften"@de;
       isoprops:countryOfOrigin "DE" ;
       isoprops:countryOfUse "DE" ;
       isoprops:creatorsLanguage "de-DE" ;
.
```

Algorithm 1: Turtle template illustrating the structure of a GroupOfProperties instance.

The ontology itself comprises two main classes, namely *Groups of Properties*, which represents collections of related properties, and *Property*, which denotes individual, measurable, or descriptive

features of an Item. A template for defining a *Group of Properties* instance in Turtle syntax is provided in Algorithm 1. As illustrated, each GroupOfProperties is further classified using the object property `isoprops:categoryOfGroupOfProperties` into one of five predefined categories. These categories are *Domain*, *Alternative Use*, *Composed Property*, *Class*, or *Reference Document*, all of which are explicitly defined in the prompt. In addition, each group includes semantic metadata such as its name, the language of origin, country of use, and, where not explicitly provided in the source text, a marked AI-generated definition. For individual properties, the ontology captures metadata like measurement units, associated physical quantities, value boundaries, and precision. The ontology also supports enumerated values, boundary conditions, and measurement methods, where applicable. Each element is linked to its corresponding Group of Properties. Further key modeling elements are those that capture domain-specific constraints such as measurement precision, value ranges, and unit definitions. Rather than modeling these details separately, the methodology incorporates them directly into the prompt structure through clearly defined instructions and few-shot examples.

4 Evaluation

To evaluate the proposed approach, a set of exemplary regulatory texts was selected from publicly available building guidelines. These texts were carefully curated to cover a range of typical information requirements, including both explicit and implicit constraints. The size and complexity of the input documents were deliberately constrained to remain within the token limit of the Mistral API (approximately 32,000 tokens), ensuring reliable model performance and response consistency.

The evaluation addressed the two main aspects under consideration, namely the quality of the extracted requirements and the correctness of the generated ontology. Firstly, the LLM was prompted using tailored templates designed to elicit both specific requirement statements and contextual interpretations. A qualitative human-in-the-loop review assesses the technical accuracy and domain relevance of the results, with experts in building regulations and semantic modeling verifying that the generated requirements and ontology entries accurately reflected the intended regulatory constraints. For the second aspect, the extracted information was instantiated as an A-Box ontology conforming to the ISOProps Ontology. The generated ontology is validated using SHACL (Shapes Constraint Language) shapes derived from the corresponding T-Box ontology, ensuring both syntactic structure and semantic alignment with ISO 23386. Any violations or inconsistencies detected during the SHACL validation were analyzed to refine the prompt design or improve the mapping logic. After iterative refinement of the prompt design, the LLM ultimately produced ontology instances with correct syntax that complied with the defined SHACL shapes.

Overall, the evaluation demonstrates the feasibility of using LLMs to automate the extraction and structuring of building information requirements. While effective for small to moderately sized documents, performance scalability and accuracy across broader and more diverse regulatory corpora remain areas for future investigation.

5 Conclusion

This paper presented an automated method for extracting and structuring information requirements from regulatory building texts using an LLM combined with prompt engineering techniques. By

translating unstructured regulatory content into a machine-readable A-Box ontology aligned with the ISOProps Ontology and ISO 23386, the approach supports greater automation and consistency in digital building requirement validation workflows. The evaluation demonstrated that the method is capable of identifying both explicit and implicit requirements and encoding them in a semantically compliant format. Furthermore, the use of SHACL shapes for ontology validation ensures formal alignment with the defined T-Box structures[3].

However, the approach is not without limitations. One key constraint is the token limit imposed by the Mistral API, which restricts the size of input documents that can be processed in a single inference. This necessitates either careful document segmentation or future strategies for handling larger inputs in an iterative manner. While LLMs demonstrate significant capabilities in processing and interpreting natural language, they are not without limitations. Errors such as inaccurate information extraction, contextual misunderstandings, or improper mappings may arise, especially when dealing with ambiguous or complex regulatory language. Consequently, incorporating a human-in-the-loop is crucial to ensure the accuracy and relevance of the extracted information.

Despite these drawbacks, the proposed approach offers a promising direction for semi-automated digitization of building regulations. Future work will focus on extending to larger document sets, establishing a quantitative evaluation metric, refining prompt strategies to improve accuracy, and exploring hybrid techniques that combine rule-based and LLM-based extraction to enhance robustness.

References

[1] S. Fuchs, J. Dimyadi, M. Witbrock, and Amor Robert, “Improving the semantic parsing of building regulations through intermediate representations”, *The 30th EG-ICE: International Conference on Intelligent Computing in Engineering*, 2023.

[2] Z. Zhang, L. Ma, and T. Broyd, “Towards fully-automated code compliance checking of building regulations: Challenges for rule interpretation and representation”, *European Conference on Computing in Construction EC3*, 2022. DOI: 10.35490/EC3.2022.148

[3] Y.-C. Zhou, Z. Zheng, J.-R. Lin, and X.-Z. Lu, “Integrating nlp and context-free grammar for complex rule interpretation towards automated compliance checking”, *Computers in Industry*, vol. 142, p. 103 746, 2022. DOI: 10.1016/j.compind.2022.103746 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166361522001439

[4] M. M. Filardo, L. Liu, P. Hagedorn, S. Zentgraf, J. Melzner, and M. König, “A standard-based ontology network for information requirements in digital construction projects”, in *Proceedings of the 12th Linked Data in Architecture and Construction Workshop, Bochum, Germany, June 13-14, 2024*, P. Pauwels, M. Poveda-Villalón, and W. Terkaj, editors, ser. CEUR Workshop Proceedings, vol. 3824, CEUR-WS.org, 2024, pp. 77–90. [Online]. Available: https://ceur-ws.org/Vol-3824/paper6.pdf

[5] S. Zentgraf, P. Hagedorn, and M. König, “Multi-requirements ontology engineering for automated processing of document-based building codes to linked building data properties”, *IOP Conference Series: Earth and Environmental Science*, vol. 1101, no. 9, p. 092 007, 2022. DOI: 10.1088/1755-

[3] Relevant T-box ontologies available at: https://github.com/RUB-Informatik-im-Bauwesen/ir-ontologies

1315/1101/9/092007 [Online]. Available: https://iopscience.iop.org/article/10.1088/1755-1315/1101/9/092007

[6] W. Solihin, Z. Liu, Y. Lu, and L. Wei, "Bim-based automated rule-checking in the aeco industry: Learning from semiconductor manufacturing", *Automation in Construction*, vol. 162, p. 105 406, 2024. DOI: https://doi.org/10.1016/j.autcon.2024.105406 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580524001420

[7] J. Zhang, B. Cui, Y. Gao, and D. Zhang, "Factors influencing the acceptance of bim-based automated code compliance checking in the aec industry in china", *Journal of Management in Engineering*, vol. 39, no. 6, p. 04 023 036, 2023. DOI: 10.1061/JMENEA.MEENG-5344

[8] E. Nuyts, M. Bonduel, and R. Verstraeten, "Comparative analysis of approaches for automated compliance checking of construction data", *Advanced Engineering Informatics*, vol. 60, p. 102 443, 2024. DOI: https://doi.org/10.1016/j.aei.2024.102443 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1474034624000910

[9] P. Pauwels, E. van den Bersselaar, and L. Verhelst, "Validation of technical requirements for a bim model using semantic web technologies", *Advanced Engineering Informatics*, vol. 60, p. 102 426, 2024. DOI: 10.1016/j.aei.2024.102426 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1474034624000740

[10] *Publikationen - Richtlinien für den Entwurf, die konstruktive Ausbildung und Ausstattung von Ingenieurbauten (RE-ING)*, 30.04.2025. [Online]. Available: https://www.bast.de/DE/Publikationen/Regelwerke/Ingenieurbau/Entwurf/RE-ING.html

[11] G. Sasson Lazovsky, T. Raz, and Y. N. Kenett, "The art of creative inquiry—from question asking to prompt engineering", *The Journal of Creative Behavior*, vol. 59, no. 1, 2025. DOI: 10.1002/jocb.671

[12] S. Shahriar et al., *Putting gpt-4o to the sword: A comprehensive evaluation of language, vision, speech, and multimodal proficiency*, 2024. arXiv: 2407.09519 [cs.AI]. [Online]. Available: https://arxiv.org/abs/2407.09519

[13] H.-C. Tsai, Y.-F. Huang, and C.-W. Kuo, *Comparative Analysis of Automatic Literature Review Using Mistral Large Language Model and Human Reviewers*. 2024. DOI: 10.21203/rs.3.rs-4022248/v1 [Online]. Available: https://assets-eu.researchsquare.com/files/rs-4022248/v1_covered_31b35741-397a-47a8-aa52-d71a4150153f.pdf

[14] S. Schulhoff et al., *The prompt report: A systematic survey of prompt engineering techniques*. [Online]. Available: http://arxiv.org/pdf/2406.06608v6

[15] L. Höltgen, S. Zentgraf, P. Hagedorn, and M. König, "Utilizing large language models for semantic enrichment of infrastructure condition data: A comparative study of gpt and llama models", *AI in Civil Engineering*, vol. 4, no. 1, 2025. DOI: 10.1007/s43503-025-00055-9

[16] O. Iversen, "Leveraging large language models for bim-based automated compliance checking of building regulations", Ph.D. dissertation, NTNU. [Online]. Available: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/3149999

An ontology-based approach for sewer infrastructure management

Raquel Valles Gómez[1], Luis Jamora[1], Peter Gölzhäuser[1], Sabine Hartmann[1], Amirhossein Babaei Ravandi[2], Hendrik Morgenstern[1] and Katharina Klemt-Albert[1]

[1]Institute for Construction Management, Digital Engineering and Robotics in Construction, RWTH Aachen University

[2]Intelligent Buildings Digital Twinning, Dept. of Civil & Mineral Engineering, University of Toronto

E-mail(s): valles@icom.rwth-aachen.de, l.jamora@icom.rwth-aachen.de

Abstract: Underground sewer infrastructure management faces challenges due to fragmented, analog asset data. While digital twins offer promising solutions for real-time monitoring and predictive maintenance, their implementation comes with challenges in data integration. The German DWA-M 150 standard provides an established basis for data exchange, yet its Extensible Markup Language (XML) schema offers limited semantic depth and lacks support for automated reasoning. We propose M150-Onto, a new domain-specific ontology that extends the DWA-M 150 schema with formal semantics. The research begins with a review of existing ontologies for sewer systems, revealing that current solutions do not fully address the requirements of German data standards for asset management. By enriching the XML schema with formal semantics, M150-Onto enables automated reasoning, complex queries, and the integration of asset, geometry, and condition data into future digital twin environments. Validation with real municipal datasets confirms its ability to support analytical tasks that are impractical with XML alone. This approach contributes to a data-driven decision-making process in the operation of sewer networks, allowing proactive maintenance and improving network resilience. Future work will extend the ontology to incorporate sensor data and damage classification.

Keywords: Ontology, Building Information Modeling (BIM), sewer infrastructure, operation and maintenance, digital twin

DOI: 10.18154/RWTH-CONV-254905. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Germany's sewage infrastructure is a dynamic system, both aging and expanding with new constructions [1]. A 2020 DWA survey [1] reports a steady rise in annual pipeline construction. The average system age is 36.9 years, with 65% of pipelines aged 0–50 years, indicating both expansion needs and modernization demands. Financial constraints and limited political support frequently lead to reactive rather than proactive maintenance [2], while climate impacts further strain the aging infrastructure, overwhelming system capacities and necessitating adaptive strategies [3]. However, a primary difficulty for the adoption of predictive maintenance methods lies in managing and exchanging large amounts

of data, modeling component relationships, and formalizing operational knowledge including physical asset details and real-time operational metrics. Traditional methods struggle to handle this diversity and volume efficiently [4] [5].

Digital technologies, mainly Geographic Information Systems (GIS), have become an established solution for spatial analysis and visualization of infrastructure data [6]. Sewer information systems ("Kanalinformationssysteme", or KIS) are specialized GIS-based applications designed for the collection, documentation, maintenance, visualization and analysis, as well as processing and exchange of sewer infrastructure data [7]. Building Information Modeling (BIM) is gaining importance in sewer infrastructure management, providing 3D visualizations of sewer components with detailed geometric and semantic information and enabling efficient information retrieval, querying, and updates [8]. Integrating data from these heterogeneous sources – GIS, KIS, and BIM – presents significant challenges due to distinct data structures, semantics, and exchange protocols. Current data exchange relies heavily on XML-based formats, particularly standardized schemas like DWA-M 150, which ensure structural validation but lack formal semantics. Ontologies address this gap by adding semantic richness and reasoning capabilities, enabling integration, inference, and advanced queries across diverse datasets. As a result, ontology engineering provides a unified framework for representing sewer infrastructure data, supporting advanced analysis and informed decision-making [9].

1.1 Background

Sewage systems, as highlighted by the DWA [10], can be described using a node-edge model in which nodes represent structural elements – such as manholes, water treatment plants, and inspection chambers – while edges denote the connecting pipe sections. Each of these elements has attribute data essential for operation, maintenance, and planning. For instance, manholes (nodes) include dimensional and structural properties, while pipes (edges) are described by material, diameter, gradient, and installation date. In Germany, the management and standardization of this data are governed by frameworks such as the DWA-M 150 guideline [11] and the federal ISYBAU standard [12].

The DWA-M 150 guideline also sets out detailed requirements for the underlying data model and its structure, based on XML and XML Schema Definition (XSD) architectures. Data exchange within sewer information systems (KIS) is implemented through XML files that must be validated against the appropriate XSD schemas, which define and constrain the structure and classification of each data element.

Table 1: *Example of identifiers for "Pipe Section: Basic data" (HG) in DWA-M 150*

Identifier	Field Name	Description
HG001	Pipe Section Designation	According to project specifications
HG002	Alternative Designation	e.g., from historical designation
...	...	...
HG006	Object Type	Type from which branching occurs (see Ref. Table 128)

Each infrastructure element – such as pipe sections ("Haltungen", HG) and nodes ("Knoten", KG) – is described by a set of standardized identifiers, which are defined as mandatory or optional within

the schema. For instance, for a "Pipe Section" element, the basic data category is described with the identifier HG, as shown in Table 1. The XSD schema moreover defines four format types (A, B, D, Z) to cover various data exchange scenarios – ranging from the transfer of master data to inspection results and assessments. Format Type Z, which is the template for individual information, encompasses all elements in the leaflet, regardless of user type or intended usage, covering all the intended use cases for data integration. Altogether, these frameworks and standards form the foundation for the approach presented in this work, which comprises four main phases: (1) requirements analysis and assessment of existing ontologies using a structured comparison; (2) concept development including ontological modeling and metamodel definition; (3) implementation of the ontology structure, axioms, and automated data migration tools; and (4) validation using real-world and synthetic data for population and query testing.

2 Related Work

A literature review was conducted to identify, evaluate, and synthesize relevant research on ontologies applicable to sewer and pipe segments. The search utilized a combination of three primary terms: Ontology, Sewage System, and Data – each accompanied by a set of alternative terms. The results were refined to ensure relevance: First, the publication date was prioritized, with a focus on articles published between 2015 and 2025. Second, relevance was strictly maintained by concentrating on studies directly related to sewer networks, computer science, and engineering. Utility networks unrelated to these areas were excluded, unless they presented methodologies that could be directly transferred and applied to the research scope. It is worth noting that articles within wastewater treatment plants were also prevalent within the search results, though this area falls outside the primary scope of this work. Following a comprehensive review of available semantic models in the infrastructure domain, three ontologies and a subontology have been identified for closer analysis and investigation. These ontologies were selected based on their relevance to underground infrastructure, their ability to represent physical asset properties, and their potential compatibility with condition assessment procedures mandated by German sewage infrastructure standards.

The ATU ontology by Du et al. [5] serves as a semantic framework for buried infrastructure – such as pipes, roads, and cables – supporting integrated urban asset management. Its Water Pipe Ontology is a specialized subontology focused on water systems, detailing pipe properties, materials, defects, and deterioration classifications. OGC GeoSPARQL, introduced by Car et al. [13], is a standard for semantic geospatial data, enabling spatial relationships, geometry representation, and infrastructure mapping through advanced querying. The Cats_Onto ontology by Zeb et al. [14] provides decision support for condition assessment technologies by linking pipe assets with their properties, defects, and inspection classifications. Additionally, the Utility Operation and Maintenance Ontology (UOMO) by Wang [15] was considered, but ultimately excluded due to the lack of open-access files and its limited validation using U.S.-based data only.

The analysis and evaluation of existing ontologies followed a structured framework based on their ability to represent the element classifications from DWA-M 150. For each ontology, we counted how many DWA-M 150 elements could be implemented directly (i.e., mapped without extensions), thereby assessing coverage of the German sewage network data structures. The evaluation categories

directly correspond to the elements outlined in DWA-M 150, including basic, geometry, inspection, and condition data and organized into two primary domains: Pipe Sections (Haltungen) and Nodes (Knoten), as shown in Table 2.

Table 2: Overview of element categories and coding according to DWA-M 150

Element	Basic Data	Geometry Data	Inspection Data	Condition Data
Pipe Section	HG	GO/GP	HI/HM	HZ
Nodes	KG/KA	GO/GP	KI	KZ

A critical insight from this evaluation is that these ontologies fail to integrate both node and pipe section data. Given the absence or limited representation of node-related data, the analysis concentrated exclusively on pipe section data across all ontologies. While some, such as ATU/Water Pipe Ontology and OGC GeoSPARQL, demonstrate strength in specific categories (e.g., Geometry Data), none offer comprehensive support for all four pipe section data categories. Furthermore, the consistent lack of node data integration – encompassing basic/structural data (KG/KA), geometry data (GO/GP), inspection data (KI), and condition data (KZ) – restricts their utility for holistic sewage infrastructure management.

3 Concept

Based on the insights gained from the literature review and the evaluation of existing ontologies, a new ontology was developed. This ontology aims to implement the data elements described in DWA-M 150 within a formal knowledge domain. The ontology development follows the methodological approach as described in [14]. The intended use of M150-Onto is to facilitate knowledge queries for predictive maintenance decision support through automated reasoning and enhanced querying for operational decisions, including the following use cases: inventory management (basic and geometry data), condition monitoring and maintenance planning (inspection and condition data), and future data integration from heterogeneous sources (all data categories) into a digital twin environment. The scope is limited to modeling network infrastructure components (e.g., pipes, nodes), excluding treatment processes. Target users include professionals in maintenance and planning roles.

For the metamodel definition, a spreadsheet was developed to manually map every element found in the DWA-M 150 advisory leaflet to suitable Web Ontology Language (OWL 2) ontology constructs. A snippet of the spreadsheet including basic data for pipe elements (analogous to Table 1) is shown in Figure 1. The analysis revealed that the majority of data elements could be represented as OWL data properties, with formats specified in DWA-M 150 Type Z (e.g., `xsd:float` for average depth). Initially, subontologies were organized by data category, but analysis showed significant overlap between elements for pipes and nodes. The revised model consolidates shared elements (e.g., designations) under common data properties (e.g., `hasDesignation`), while unique elements are grouped by functional relevance within the data hierarchy.

Element	English Translation	OWL Component	Label (@en)	Format
HG/KG	**Basic Data**	**(Super) Data Property**	**hasBasicData**	**xsd:boolean**
HG001	Pipe Section Designation	Data Property	hasDesignation	xsd:string
KG001	Node Designation			
HG002	Alternative Pipe Section Designation		hasAlternativeDesignation	xsd:string
KG002	Alternative Node Designation			

Figure 1: Snippet of the metamodel for basic data

M150-Onto comprises four main taxonomies: The component taxonomy includes the main physical elements of the system, such as pipe sections and nodes. The geometry taxonomy encompasses the GIS-based geometric information associated with these components, including both geometry objects and geometry points that represent their spatial properties. The report taxonomy structures the various reports related to system components, such as inspection and condition reports, in alignment with the formats specified in DWA-M 150. Finally, the reference taxonomy incorporates the standardized reference tables provided in the advisory leaflet, covering material types, coding systems, and other classifications.

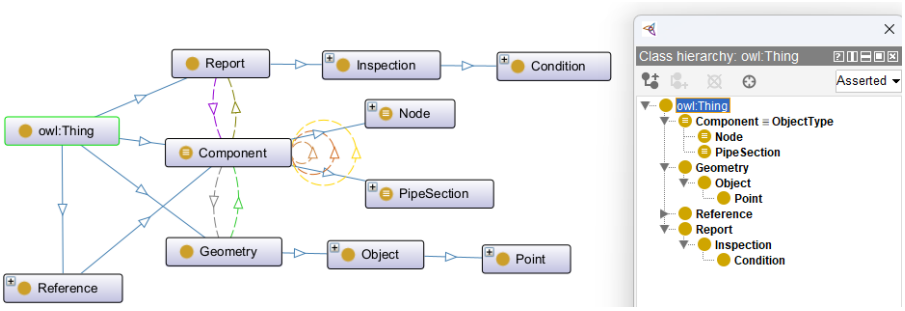


Figure 2: Class Hierarchy

4 Implementation

Ontology coding was carried out in Protégé, translating the previously defined metamodel and taxonomies into OWL. Each taxonomy became a superclass with subclasses mirroring the M150 structure, as shown in Figure 2. Component includes `PipeSection` and `Node`, with flow relationships described by the object property `flowsTo` (and its inverse, `flowsFrom`), and inferred network connections via `connectedWith`. Geometry is modeled with `Object` and its subclass `Point`, reflecting the containment of geometry points within geometry objects, using `isParentOf` and `renders` properties. The report taxonomy includes `Inspection` and its subclass `Condition`, aligned with DWA-M 150's structure for inspection and condition data. Reference tables from DWA-M 150 are represented as subclasses (e.g., `Material`, `CodingSystem`) linked to their respective table numbers. Equivalent class axioms (`owl:equivalentClass`) link ontology classes to reference table codes (e.g., `Material hasReferenceTable value "105"`). Specific subclasses, such as `Concrete`, are defined by their material codes (e.g., `hasMaterial value "B"`). Disjoint axioms (`owl:disjointWith`) ensure logical separation between classes that cannot overlap. In M150-Onto, for example, the class `Node` is disjoint from the class `PipeSection`, indicating that a component cannot simultaneously be both a node and a

pipe section, despite both being subclasses of `Component`. Additionally, the axiom `rdfs:subClassOf` defines hierarchical relationships between classes.

4.1 Mapping of Inventory Data to M150-Onto

The process of manually importing individual pipe sections and nodes was found to be highly time-consuming. An automated import tool for translating M150-XML into OWL individuals based on a Python script was developed to enable the integration of existing inventory data into M150-Onto. The script imports the M150-Onto ontology using the `owlready2` library, reads each XML element from files structured according to DWA-M 150, identifies the corresponding ontology classes and properties based on the DWA-M 150 schema, and finally outputs M150-Onto with the new individuals.

5 Testing and Validation

To demonstrate the practical application of M150-Onto, sample data from a real municipal sewage system in Germany was populated as individuals within the ontology. The data was originally provided in M150-XML format and included basic/structural data (HG/KG) and geometry data (GO/GP) for 701 pipe elements and 702 node elements. Since the data did not include inspection data (HI/KI) and condition data (HZ/KZ), synthetic inspection and condition data were created for a pipe section to test the ontology, as shown in Figure 3.

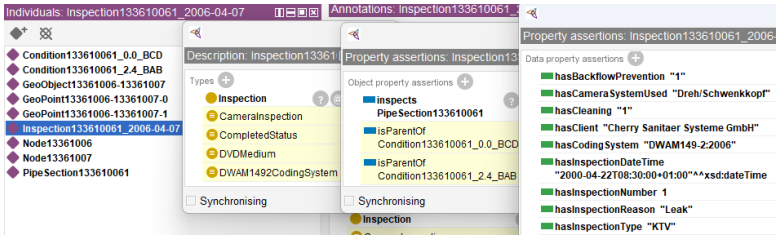


Figure 3: Example of synthetic inspection data (HI) for a pipe section

Ontology consistency was verified using the Pellet reasoner integrated with Protégé. The reasoner identified inconsistencies in the real-world data due to three pipe sections having the same designation. After correction, a series of queries was applied to the data for validation: identifying all combined sewers within a district, locating concrete pipe sections with cracks, retrieving specific condition reports related to pipe sections, and inferring the connections between pipe sections and nodes. All queries returned correct and expected results. Additionally, the ontology was tested with a sample dataset from DWA (all data categories) using the same consistency check and queries for validation.

6 Conclusion and Outlook

The development of M150-Onto was generally unproblematic due to the clear class-and-element structure of the XML files, which closely matched the ontology's conceptual model. However, manually mapping the M150-XML data structure to the ontology during development was slow and demanding. While M150-XML data could partially be implemented within some of the reviewed existing ontologies after minor accommodations, the overall mapping process – including later ontology population and linking the data – would have been much more difficult and time-consuming. Therefore, creating a tailored ontology (M150-Onto), designed specifically for the current M150-XML format, ultimately

reduces complexity. Future adjustments to M150-Onto can be made to align with the new DWA-M 145-3 data structure (currently in draft). For ontology population, an automated mapping process was developed to import real data, mapping the XML elements to OWL. So far, this automated import has only been tested with two dataset, due to limited data availability. Challenges may arise when processing larger or more complex inventories.

Several challenges and limitations remain. The ontology has not yet been validated by domain experts, limiting its reliability for operational use. Its design is based on German standards, which may restrict its applicability in other contexts. While some DIN regulations serve as a common framework for several EU member states, and thus offer potential for broader adoption, application beyond such regulatory environments may prove difficult. Another key limitation is the current lack of implementation of damage classifications based on DIN EN 13508-2, which defines a comprehensive coding system for condition assessment of drain and sewer systems, and the lack of integrated sensor data. Introducing taxonomies for damage coding systems would establish equivalence relationships between different classification methods and enable inference-based queries that support maintenance decision-making. Future work will extend the ontology to cover these elements and develop a data layer that supports automated asset data processing across diverse sources. This will allow the framework to link data from sewer management systems, GIS, and BIM.

Code Availability

M150-Onto and the XML parser are available at https://l-jamora.github.io/m150-onto/ and https://github.com/l-jamora/m150-onto. M150-Onto was created in the context of Luis Jamora's bachelor's thesis on the investigation of existing ontologies for underground infrastructure.

Acknowledgements

This research is part of the "KaSyTwin" project funded by the Federal Ministry for Digital and Transport (BMDV) within the mFUND program under grant 01FV2059A.

References

[1] DWA, "Zustand der Kanalisation in Deutschland", *Zustand der Kanalisation in Deutschland*, vol. 2020, no. 12, pp. 939–953, Dec. 2020. DOI: 10.3242/kae2020.12.001

[2] J. H. Trapp, H. Kerber, and E. Schramm, "Implementation and diffusion of innovative water infrastructures: Obstacles, stakeholder networks and strategic opportunities for utilities", *Environ Earth Sci*, vol. 76, no. 4, p. 154, Feb. 2017. DOI: 10.1007/s12665-017-6461-8

[3] G. Blöschl et al., "Changing climate both increases and decreases European river floods", *Nature*, vol. 573, no. 7772, pp. 108–111, Sep. 2019. DOI: 10.1038/s41586-019-1495-6

[4] J. H. Tse, "Challenges of Adopting Knowledge-based Building Information Modeling for E&M Asset Management Supplemented with Mobile Solutions - A Case Study in Public Sewage Pumping Facilities.", *American Journal of Operations Management and Information Systems*, vol. 1, pp. 17–33, Nov. 2016. DOI: 10.11648/j.ajomis.20160101.13

[5] H. Du et al., "City infrastructure ontologies", *Computers, Environment and Urban Systems*, vol. 104, p. 101 991, Sep. 2023. DOI: 10.1016/j.compenvurbsys.2023.101991

[6] R. Attah, B. Garba, I. Gil-Ozoudeh, and O. Iwuanyanwu, “Leveraging geographic information systems and data analytics for enhanced public sector decision-making and urban planning”, *Magna Scientia Advanced Research and Reviews*, vol. 12, pp. 152–163, Nov. 2024. DOI: 10.30574/msarr.2024.12.2.0191

[7] DWA, “Kanalinformationssysteme. 1: Grundlagen und systemtechnische Anforderungen”, in ser. DWA-Regelwerk Merkblatt DWA-M 145-1, [Stand:] Dezember 2013, Hennef: Dt. Vereinigung für Wasserwirtschaft, Abwasser und Abfall, 2013.

[8] M. Wang, Y. Deng, J. Won, and J. C. Cheng, “An integrated underground utility management and decision support based on bim and gis”, *Automation in Construction*, vol. 107, p. 102 931, 2019. DOI: 10.1016/j.autcon.2019.102931

[9] T. R. Gruber, “A translation approach to portable ontology specifications”, *Knowledge Acquisition*, vol. 5, no. 2, pp. 199–220, Jun. 1993. DOI: 10.1006/knac.1993.1008

[10] DWA, editor, *Anforderungen an ein Datenmodell und Schnittstelle (Entwurf)* (Kanalinformationssysteme / DWA, Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall Teil 3), Entwurf, September 2021, 1. Auflage. Hennef (Sieg): Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall, 2021.

[11] DWA, editor, *Datenaustauschformat für die Zustandserfassung von Entwässerungssystemen* (DWA-Regelwerk Merkblatt DWA-M 150), [Stand] April 2010. Hennef: DWA, 2010.

[12] BMJ, *Gesetz zur Ordnung des Wasserhaushalts (Wasserhaushaltsgesetz)*, 2009. Accessed: Oct. 5, 2024. [Online]. Available: https://www.gesetze-im-internet.de/whg_2009/

[13] Nicholas J. Car et al., *OGC GeoSPARQL - A Geographic Query Language for RDF Data*, 2024. Accessed: Mar. 25, 2025. [Online]. Available: https://docs.ogc.org/is/22-047r1/22-047r1.html#_b93d0f57-5812-4f37-b17a-07cc54e2bae6

[14] J. Zeb, “An ontology of condition assessment technologies for sewer networks”, *Infrastructure Asset Management*, vol. 7, no. 1, pp. 36–45, Mar. 2020. DOI: 10.1680/jinam.18.00034

[15] M. Wang, “Ontology-based modelling of lifecycle underground utility information to support operation and maintenance”, *Automation in Construction*, vol. 132, p. 103 933, Dec. 2021. DOI: 10.1016/j.autcon.2021.103933

Digital Heritage in VR: 360° Media, Theological Insights and Interactive Artifact Exploration – The Case of Tel Lachish

Djamel Berkaoui[1], Christina Kumpmann[2], Koen Castermans[1], Raymond Leonardo Chandra[1], Patrick Querl[1] and Heribert Nacken[1]

[1]Academic and Research Department Engineering Hydrology and UNESCO Chair of Hydrological Change and Water Resources Management, RWTH Aachen University, Aachen, Germany

[2]Catholic University of Applied Sciences, Mainz, Germany

E-mail(s): berkaoui@lfi.rwth-aachen.de, christina.kumpmann@kh-mz.de, castermans@lfi.rwth-aachen.de, chandra@lfi.rwth-aachen.de, querl@lfi.rwth-aachen.de, nacken@lfi.rwth-aachen.de

Abstract: The usage of Virtual Reality (VR) in higher education opens up new possibilities to enhance teaching to be more interactive and immersive. The following study examines the development of a use-case for theology built on the MyScore open-source VR framework rooted in engineering hydrology, a VR application extending to construction informatics by enabling digital twins for historical site reconstruction and sustainable analysis, reducing carbon footprints from physical excursions. The application enables multi-user, location- and time-independent interaction in VR with the historical site of Tel Lachish. It combines 360° media with a digital learning space that provides in-depth historical content, as well as an interactive environment featuring a puzzle-based reconstruction of a museum relief, allowing for exploratory engagement with ancient artifacts. The use of mixed-method approaches, combining curated questionnaires and in-situ observations, shows how the VR environment supports understanding of historical and theological contexts and explores the associated challenges. The findings show that the immersive combination of virtual environments and interactive learning elements not only increases interest in the subject matter, but also enables new forms of knowledge acquisition. The possibility to learn in a self-paced and mentor-guided manner was perceived as flexible and effective by the users. However, technical issues, especially in terms of device usage and stable internet connectivity, were noted. This study highlights the potential of using VR to enhance traditional teaching methods by creating location-independent, collaborative and interactive learning environments, which are tailored as a support for higher education, where physical access to historical sites is limited.

Keywords: Virtual Reality, Theology, Unity, Immersive Learning, Higher Education

DOI: 10.18154/RWTH-CONV-254906. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

1.1 Problem Statement

In theological education, historical sites such as Tel Lachish often remain abstract, since they are usually conveyed through texts or two-dimensional media [1]. Historical-critical biblical science requires contextualization of ancient texts, typically deepened through real-world excursions a method hindered by costs, access restrictions and logistical challenges [2]. VR provides innovative ways to make historical contexts accessible and immersive by integrating archaeological findings with textual narratives [3]. Using the example of the Assyrian conquest of Lachish (701 BC), a key archaeological site for understanding biblical narratives, this study examines how VR-supported learning environments promote theological education by deepening historical understanding and commitment, as well as enabling historical building preservation and sustainable site analysis. The application was developed within the open-source VR framework MyScore [4], targeting theology students. The framework, initially tailored, for civil engineering scenarios like flood protection and virtual site modeling, bridges theological heritage with informatics tools for building preservation and digital sustainability [4].

1.2 Objective and Research Question

The goal of this study is to examine the didactic potential of VR and the provided learning environment for historical-theological content. The main focus is to analyze how VR enhances motivation, historical empathy and comprehension through the implementation of traditional media (text, images, videos) in an immersive environment, as well as using constructivist principles [5] for an interactive teaching environment. Research questions:

1. How does VR foster motivation, empathy and contextual comprehension in historical-theological learning?
2. What learning effects arise from combining classical media with immersive spatial experiences?
3. What potential and limits does VR have as an excursion alternative?

The use-case was designed primarily for theology undergraduates to contextualize biblical sites, whereas MyScore's VR framework makes it adaptable for construction-related domains like hydrology engineering education (e.g., virtual flood simulations) and digital heritage preservation in building informatics [4]. This study emphasizes VR in education with sustainability via MyScore's alignment to SDG 4. While VR has been widely applied in engineering, medicine and language learning [6], [7], applications in theological or religious-historical contexts remain underrepresented. Recent systematic reviews [6], [7] list few or no examples that combine archaeological reconstruction, theological interpretation and interactive artifact engagement, marking this as a niche with distinct didactic potential.

2 Theoretical Framework

2.1 Historical-Critical Bible Science and Digital Contextualization

The historical-critical exegesis examines biblical texts as literary testimonies in their historical and cultural context [8]. VR enhances textual work with the addition of visual and spatial components,

enabling historical empathy and illustrating the difference between text and reconstructed history [6], [9], [10]. VR visualizes archaeological textual discoveries, enabling learning with multi-perspectives and building a bridge to abstract text analysis with historical contexts [6].

2.2 Virtual Reality in the Educational Context

VR offers high potential in didactics with the use of immersion, presence and interactivity [11]. Immersion enables diving deep into the virtual environment, presence creates a feeling of attendance with avatars and interactivity promotes active exploration [12], [13]. Incorporation of 360° media enables a change of perspective, while playful elements such as puzzles increase motivation and cognitive activation [14]. Multi-user functions support social interactions and communication in collaborative learning environments [12], [13]. Immersive media creates emotional approaches to promote historical empathy [6]. In theological education, real-world excursions are a core component of methodological competence and VR offers a scalable, accessible alternative to physical visits (e.g. Lachish or Jerusalem) [15]. Gamified methods increase participation through playful, problem-solving-oriented tasks that are coordinated with learning objectives [5], [7].

3 System Development

3.1 Technological Basics

The VR use-case of Tel Lachish was created with the help of Unity Engine (2022 LTS) and the open-source VR framework MyScore [4]. Additionally, Blender for 3D modelling was incorporated in the development process. The Meta Quest 2 and Meta Quest 3 standalone VR devices were used for the testings. The VR framework enables the integration of 360° media, interactive actions and multi-user functionalities in a modular design. It is optimized for mobile headsets as well as an non-immersive Windows version. The framework ensures scalability and synchronization in multi-user environments.

3.2 Conception of Content: Tel Lachish as a Place of Learning

The VR environment contains four functionally structured areas, each containing elements of immersion, exploration, cognitive activation and cooperative interaction through intuitive interfaces, direct feedback and adaptive media control (Overview: [16]). The rooms are:

- **Meeting Room**: Starting center point with navigation aids, a meeting area for synchronous sessions with up to 20 users and a screen-sharing zone for collaborative parts of web content. The open design offers a clear overview of all modules and makes navigation easier (Figure 1).
- **360° Excursion Room**: Provides 360° media with an interactive map and laser pointer, using ‘Star’ shaped UI elements for user-friendly navigation (Figure 2).
- **Media Room**: Offers classic educational media formats with curated choices, that support the learning content of the environment and provide additional information (Figure 3).
- **Relief Room**: Enables the reconstruction of the Assyrian Tel Lachish Relief (British Museum) with 3D fragments, triggers for contextual videos and an introduction video that supports users through the task, as well as give valuable information after a successful execution (Figure 4).

Figure 1: Meeting area

Figure 2: 360° media area

Figure 3: Media room

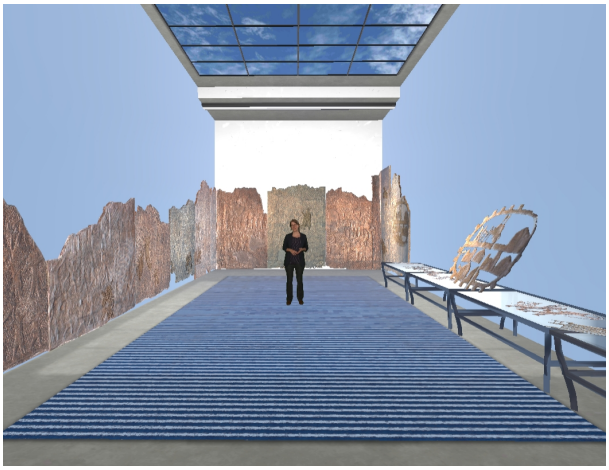


Figure 4: Relief room

3.3 Usability and Technological Challenges

For efficient use of VR devices, a well-structured introduction to the control and navigation system is needed to provide users with different technical backgrounds with the most comfortable experience. Challenges include motion sickness, limited battery life, high graphics processing demands and stable network connectivity. VR hardware and software preparation is essential to minimize technical issues.

To ensure accessibility and comfort, the VR experience was implemented within the MyScore framework with built-in comfort features such as teleport locomotion and vignette mode to reduce peripheral motion during navigation, as well as a snap turn feature. These measures address motion sickness by limiting continuous acceleration and maintaining a stable horizon, in line with VR ergonomics research [11]. While no dedicated measurement of motion sickness was conducted, no participant reported discontinuing the session due to discomfort.

Battery life and device overheating were mitigated through a low-poly visual style, which reduces GPU load and short sessions rather than prolonged exposures. Graphical demands were further reduced through light baking and efficient asset management, minimizing real-time lighting calculations. These

methods are known to improve battery longevity without noticeably impacting immersion [9], [10] and together they supported smoother performance without compromising learning content.

4 Didactic Design

The VR use-case "Tel Lachish" was developed as an immersive, skill-oriented and multimedia learning environment for theological and historical education. The main focus was offering historical contextualization, digital skills and broad learning methods. The design corresponds to the concepts of constructive alignment, aligning objectives, immersive tasks, and assessment opportunities.

4.1 Learning Objectives

The learning environment supports:

- **Understanding historical contexts**: Learners are encouraged to engage with biblical contexts in a spatial and visual way, helping them to better grasp the historical settings and narratives behind the scriptures.
- **Developing methodical skills**: The VR environment supports the development of practical skills in using digital media formats such as 360° media, videos and interactive puzzle pieces to enable learners to navigate and examine content effectively.
- **Gaining digital and communicative skills**: Through collaborative multi-user activities in a shared space, students strengthen their ability to use modern learning technologies and interact meaningfully with others in the virtual environment.

4.2 Learning Modes and Learning Experience

A self-directed exploration for autonomous learning and a mentor-based mode with structured guidance were offered. Both showed complementary advantages in autonomous versus contextual support. The learning experience follows flow and constructivist principles, enhancing motivation, emotional engagement and knowledge gain [5], [7]. Flow enables balanced challenges and direct feedback in the virtual environment, while constructivism enables interactive tasks, reflection and collaboration. Additionally, flow improves emotional engagement and historical empathy.

4.3 Integration into the Teaching Context

The application is versatile for digital excursions, seminars or extracurricular activities and offers a low-threshold entry into biblical-historical learning. It enables the combination of experiences with traditional text analysis in workshop formats or as a cost-efficient alternative to physical study trips. The modular design allows teachers to adapt content to specific curricular needs, which increases applicability in different teaching contexts.

4.4 Evaluation Methods

A mixed-method approach combined quantitative Likert-scale questionnaires (motivation, usability, satisfaction) and qualitative in-situ observations (interaction, communication). The quantitative evaluation used a 5-point Likert scale (1 = Strongly agree, 5 = Strongly disagree), chosen for its balance between sensitivity and participant reliability in small-sample studies [5], as well as multiple-choice selections. The full questionnaires and results are available publicly via the project's OSF repository [17]. The questionnaire contained five core items:

- The VR environment improved my understanding of the historical context.
- The interactive elements increased my motivation to learn.
- The navigation was intuitive and easy to use.
- The 360° media enhanced the realism of the experience.
- I would like to use VR for similar learning activities in the future.

Participants: N = 21 (mean age = 28.9 years, SD = 8.1; range = 19–54), 14 female, 6 male, 1 divers. Participants (theology students and scientific staff) volunteered in a seminar context. Qualitative data were gathered from in-situ observation of user interaction, communication and task completion.

Descriptive statistics were used for all items; no regression or correlation analysis was performed due to the exploratory scope and small sample size. A control group was not implemented in this study but is planned for future iterations to compare with traditional methods. Standardized usability instruments such as SUS or UMUX were not applied here, which may limit generalizability, but will be integrated in future studies to enable direct comparability with other systems.

5 Results and Evaluation

The mixed-method analysis with 21 participants showed:

- 66,7% (14/21) had no VR experience, underlining the need for a guided introduction. An introductory environment supported 78,9% (16/19) to understand the VR control.
- The VR application was considered by most users as motivating and helpful, based on the available learning content (e.g., "helps to better understand content," "playful learning"). 47.6% are unsure to use VR in the future, 23.8% plan to use it, 23.8% don’t plan to use it .
- The contribution to deeper understanding of historical context with puzzle-based reconstruction was rated as 30% significant, 40% slight, 20% minimal, 10% confusing.
- 85,7% (18/21) found the 360° media to be "quite" or "very immersive" and 73,7% (14/19) rated the instructional videos as "very" or "quite effective" for additional information.
- 81% (17/21) expressed interest in applying VR experience to other topics or places in a different context (e.g., Jerusalem).
- 75% (15/20) preferred mentor-based teaching units through a lecturer.

6 Discussion

The questionnaire and observations suggest that the VR application was perceived as motivating and helpful for developing historical understanding by the participants, particularly through the use of 360° media and interactive puzzles. It supports both self-directed and guided learning. Despite technical challenges, VR remains a scalable addition to teaching. Comfort depends on smooth interaction

design, while battery limits, space (2.5 m²) and stable networks remain key for multi-user use. Given the small sample size, descriptive statistics were used to identify trends, reserving advanced tests for future controlled studies.

7 Conclusion and Outlook

The presented study explored if VR can enhance student motivation, historical empathy and contextual understanding in historical-theological learning. Immersive 360° media, puzzle-oriented artifact exploration and combination of independent and guided reflection were perceived as engaging and helpful learning strategies. Most participants reported increased motivation and improved comprehension, though these findings remain tentative due to the small sample size and reliance on descriptive measures. Despite limitations related to hardware, network availability and setup requirements, the "Tel Lachish" application illustrates how VR can function as a flexible, location-independent alternative to traditional excursions, enabling collaborative and interactive learning experiences. The combination of archaeological reconstruction, theological interpretation and interactive artifact engagement may offer a perspective less common in current VR learning contexts [9], [15]. Future research approaches should add additional analytic methods as controlled comparisons, correlation and regression tests and standardized usability measures (e.g., SUS, UMUX) to investigate the effects of VR on motivation, learning outcomes and knowledge retention in a more quantified manner. Longitudinal research with larger and more diverse samples could track learning gains and experimental designs with control groups that use traditional teaching methods, as well as enabling stronger claims about the advantages of VR. Additionally, the current VR framework can be adapted for other disciplines to expand its content base. By resolving these research deficits and technical limitations, virtual reality could position itself as a scalable alternative to higher education teaching and digital heritage conservation.

References

[1] M. Victor, "The impact of vr on historical education: An investigation into the effectiveness and efficacy of immersive learning experiences", *Int. J. Hist. Res.*, vol. 3, no. 2, pp. 27–38, 2023. DOI: 10.47604/ijhr.2018

[2] "Contribution of contextual analysis to biblical exegesis", *Cuest. Teol.*, vol. 44, no. 102, pp. 283–300, 2017. DOI: 10.18566/cueteo.v44n102.a04

[3] D. N. Wijaya, H. Hariyono, W. D. Yudo, and F. Ariska, "The use of vr as a historical learning media: A systematic literature review", in *Informa*, 2024, pp. 80–86. DOI: 10.1201/9781003484318-12

[4] D. Berkaoui, K. Castermans, R. L. Chandra, P. Querl, and H. Nacken, "Myscore – avatar-based teaching and learning", 2022. DOI: 10.18154/RWTH-2023-02867

[5] G. Challco, I. I. Bittencourt, M. d. A. Reis, J. Santos, and S. Isotani, "Gamiflow: Towards a flow theory-based gamification framework for learning scenarios", in *Lect. Notes Comput. Sci.*, 2023, pp. 415–421. DOI: 10.1007/978-3-031-36336-8_65

[6] P. Rosendahl and I. Wagner, "360° videos in education – a systematic literature review on application areas and future potentials", *Educ. Inf. Technol.*, pp. 1–37, 2023. DOI: 10.1007/s10639-022-11549-9

[7] G. Lampropoulos and Kinshuk, “Virtual reality and gamification in education: A systematic review”, *Educ. Technol. Res. Dev.*, vol. 72, pp. 1691–1785, 2024. DOI: 10.1007/s11423-024-10351-3

[8] S. T. J. Smith, “Historical criticism: Methods”, in *The New Cambridge Companion to Biblical Interpretation*, I. Boxall and B. C. Gregory, editors, Cambridge: Cambridge Univ. Press, 2022, pp. 37–54.

[9] K. Betts et al., “An examination of the effects of vr training on spatial visualization and transfer of learning”, *Brain Sci.*, vol. 13, no. 6, p. 890, 2023. DOI: 10.3390/brainsci13060890

[10] G. B. Petersen and G. Makransky, “A study of how immersion and interactivity drive vr learning”, *Comput. Educ.*, vol. 179, p. 104 429, 2022. DOI: 10.1016/j.compedu.2021.104429

[11] S. Klingenberg, R. Bosse, R. E. Mayer, and G. Makransky, “Does embodiment in vr boost learning transfer? testing an immersion-interactivity framework”, *Educ. Psychol. Rev.*, vol. 36, no. 4, 2024. DOI: 10.1007/s10648-024-09956-0

[12] Q. Yuan and Q. Gao, “Being there, and being together: Avatar appearance and peer interaction in vr classrooms for video-based learning”, *Int. J. Human–Comput. Interact.*, pp. 1–21, 2023. DOI: 10.1080/10447318.2023.2189818

[13] S. Aseeri and V. Interrante, “The influence of avatar representation on interpersonal communication in virtual social environments”, *IEEE Trans. Vis. Comput. Graph.*, vol. 27, no. 5, pp. 2608–2617, 2021. DOI: 10.1109/TVCG.2021.3067783

[14] Z. Tacgin, “Unlocking the puzzle: Investigating problem-solving patterns in 2d and 3d vr environments through mixed methods analysis”, *Int. J. Human–Comput. Interact.*, pp. 1–15, 2024. DOI: 10.1080/10447318.2024.2369368

[15] U. Riegel and K. Kindermann, “Why leave the classroom? how field trips to the church affect cognitive learning outcomes”, *Learn. Instr.*, vol. 41, pp. 106–114, 2016. DOI: 10.1016/j.learninstruc.2015.10.004

[16] Academic and Research Department Engineering Hydrology, RWTH Aachen University, *Scenario: Tel lachisch excursion*, 2024. [Online]. Available: https://youtu.be/KfZvjuXUI1k

[17] Academic and Research Department Engineering Hydrology, RWTH Aachen University, *Full anonymized results of the conducted study*, 2025. [Online]. Available: https://osf.io/2763n

Datenanalysekonzept für die Anwendung von BIM in der Bauausführung von Infrastrukturprojekten

Tim Noack[1], Philip Markert[2], Björn Buuck[3], Caroline Barr[2], Al-Hakam Hamdan[3], Chongjie Kang[1] and Steffen Marx[1,2]

[1]Institut für Massivbau (IMB), TU Dresden, Dresden, Germany

[2]Marx-Krontal Partner GmbH, Weimar, Germany

[3]A+S Consult GmbH, Dresden, Germany

E-mail(s): tim.noack@tu-dresden.de, philip.markert@marxkrontal.com, bjoern.buuck@apluss.de, caroline.barr@marxkrontal.com, al-hakam.hamdan@apluss.de, chongjie.kang@tu-dresden.de, steffen.marx1@tu-dresden.de

Abstract: Building Information Modelling (BIM) wird durch verschiedene Vorgaben von Behörden im Infrastrukturbau wie beispielsweise dem Bundesministerium für Digitales und Verkehr (BMDV) bereits bei der Planung vieler Infrastrukturprojekte eingesetzt. Bei der Bauausführung hingegen kommt BIM aktuell nur in Pilotprojekten zum Einsatz. Im Zuge des Testprogramms „BIM2Build“ wurden und werden nun auch bei Projekten der DB InfraGO BIM-Methoden bei der Bauausführung eingesetzt. Durch systematische Analysen und Vergleiche von ausgewählten Projekten untereinander und mit vergleichbaren konventionellen Planungen wird eine Bewertung der Effizienz- und Qualitätssteigerung durch die Anwendung der BIM-Methodik angestrebt. Für diese Analyse wird ein Datenanalysekonzept erarbeitet, bei dem quantitative und qualitative Bewertungsparameter gefunden werden, die einen positiven oder negativen Beitrag von BIM zur Effizienz- und Qualitätssteigerung der Bauausführung objektiv messbar machen. Eine Bewertung der extrahierten Parameter erfolgt über Schulnoten. Um den Einfluss der unterschiedlichen Parameter auf das Endergebnis zu erfassen, wird zudem eine Wichtung der Bewertungsparameter vorgenommen. Durch die Kombination aus vergebenen Noten und Wichtung lassen sich die unterschiedlichen Infrastrukturprojekte objektiv und vergleichbar bewerten. Basierend auf den Erkenntnissen der Analyse werden in zukünftigen Schritten aus den analysierten Projekten Standardisierungs- und Optimierungspotenziale ermittelt und Handlungsempfehlungen zur effizienteren Anwendung von BIM bei der Bauausführung von Infrastrukturprojekten abgeleitet.

Keywords: Building information modeling (BIM), Bauausführung, Schieneninfrastruktur, Datenanalyse

DOI: 10.18154/RWTH-CONV-254907. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Einführung

Die Verbindung von dreidimensionaler Geometrie mit weiteren Informationen, wie bspw. Materialdaten oder Verknüpfungen zu Bauphasen, in Informationsmodellen hat die Methode des Building Information Modelling (BIM) zu einer der wichtigsten Technologien im modernen Bauwesen befördert. Das Potenzial der Methode ist enorm, weshalb auch deutsche Behörden vorhaben, BIM immer weiter in der Branche zu etablieren. Der „Stufenplan Digitales Bauen“ vom Bundesministerium für Verkehr und digitale Infrastruktur [1] und die davon abgeleitete VR-I-BIM Strategie der Deutsche Bahn AG (DB) [2] sind prominente Beispiele aus dem Infrastruktursektor. Um das volle Potenzial der Methode auszuschöpfen, sollte diese nicht nur in den Planungsphasen eingesetzt werden, wie es derzeit noch sehr häufig praktiziert wird. Durchgängige digitale Prozesse von der Planungsphase über die Bauphase bis zum Anlagenmanagement werden benötigt. Dafür werden allerdings auf der einen Seite das Know-How aller Beteiligten und auf der anderen Seite Standardisierungen und klare Vorgaben von den Auftraggebenden benötigt. Die DB InfraGO AG hat mit dem Ziel, beide Aspekte zu fördern, das Pilotprogramm „BIM2Build“ ins Leben gerufen [3]. In diesem sind verschiedenste Bauprojekte aus der Schieneninfrastruktur vereint, bei denen der Einsatz von BIM in allen Leistungsphasen, besonders aber bei der Bauausführung, getestet werden soll. Nachdem sich mittlerweile einige dieser Projekte in der Ausführungsphase befinden, oder diese sogar bereits abgeschlossen haben, wurde das gleichnamige mFUND-Forschungsprojekt „BIM2Build“ gestartet, in dem die gewonnenen Erfahrungen zur Optimierung und Standardisierung der Anwendung der BIM-Methodik in der Bauausführung von Infrastrukturbauwerken genutzt werden sollen.

2 Methodik

Das Forschungsprojekt verfolgt die Methodik der sogenannten DIKW-Pyramide. Diese basiert auf R. Ackoff 1989 [4] und J. Rowley 2007 [5], die beschrieben haben, wie im Bereich des Informationsmanagements aus Daten Weisheit generiert werden kann. Im Fall von BIM2Build hat die Pyramide den Aufbau nach Abbildung 1. Die hier dargestellte Pyramide besteht aus vier Stufen: Daten (en. *Data*), Informationen (*Information*), Wissen (*Knowledge*) und Weisheit (*Wisdom*).

2.1 Stufe 1: Daten

Für die erste Stufe der Pyramide werden Projektdaten von Schieneninfrastrukturprojekten gesammelt, gesichtet und ihre Qualität geprüft. Die Projekte stammen aus allen Bereichen des Eisenbahnbaus; Eisenbahnbrücken, (Tief-)Bahnhöfe, Tunnel und der Schienenoberbau wurden hier mithilfe von BIM geplant und ausgeführt. Es werden neben bereits abgeschlossenen Projekten auch in Ausführung befindliche Projekte ausgewertet. Die Daten setzen sich aus vertraglichen Dokumenten wie den Ausschreibungsunterlagen, den Auftraggeber-Informationsanforderungen (AIA) und dem BIM-Abwicklungsplan (BAP), BIM-Modellen aus den einzelnen Leistungsphasen, Mengenberechnungen und -abrechnungen, Kosten- und Terminplänen, Protokollen der Bauüberwachung und der Baubesprechungen, Bautagebüchern, Fotos der Bauausführung, Mängellisten und weiteren Dokumenten zusammen. Hinzu kommen Mitschriften von Interviews mit Projektbeteiligten wie beispielsweise aus dem BIM-Management. Dadurch lassen sich auch die Erfahrungen der Beteiligten einbeziehen.

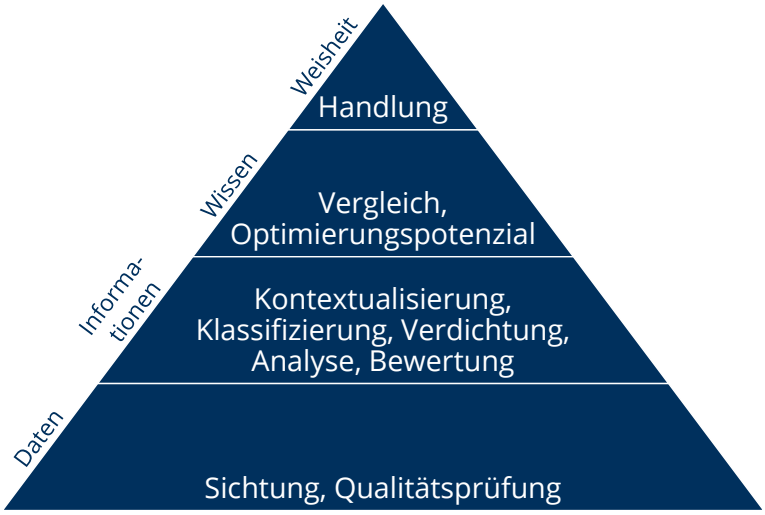


Abbildung 1: DIKW-Pyramide für die Infrastrukturprojektdaten

2.2 Stufe 2: Informationen

Dieser Schritt besteht darin, aus den Daten Informationen zu extrahieren. Das geschieht durch eine Kontextualisierung, Klassifizierung und Verdichtung, um die relevanten Daten zu ermitteln. Diese werden anschließend mithilfe eines eigens aufgestellten Datenanalysekonzepts (siehe Abschnitt 3) analysiert und bewertet.

2.3 Stufe 3: Wissen

Die Projekte werden untereinander und mit der konventionellen Planungs- und Ausführungsweise verglichen, wodurch die in Stufe 2 angehäuften Informationen vernetzt und zu Wissen umgewandelt werden. Durch die Vergleiche lassen sich Vor- und Nachteile der BIM-Methodik für die Ausführung der Infrastrukturprojekte erkennen. Ebenso lässt sich ermitteln, an welchen Stellen in der Planung und Ausführung eine Optimierung der Arbeitsweise mithilfe von BIM möglich ist.

2.4 Stufe 4: Weisheit

Die Stufe der Weisheit besteht aus der Anwendung des Wissens. Aus den in Stufe 3 ermittelten Optimierungspotenzialen lassen sich Handlungsempfehlungen zur optimalen Anwendung der BIM-Methodik ableiten. Neben den Handlungsempfehlungen werden auch Gestaltungshinweise für BIM-Modelle und Informationsaustauschprozesse in den relevanten Leistungsphasen veröffentlicht. Ziel ist eine Optimierung und Standardisierung der BIM-Anwendung für die Ausführungsphase von Infrastrukturprojekten, um durchgängige digitale Prozesse über den Lebenszyklus der betreffenden Bauwerke zu ermöglichen.

3 Datenanalysekonzept

Für die Schieneninfrastrukturprojekte aus dem Testprogramm BIM2Build der DB InfraGO wurde ein Datenanalysekonzept entwickelt, das eine objektive und vergleichbare Bewertung der Anwendung von BIM ermöglichen soll. Das Datenanalysekonzept ist der zentrale Baustein der zweiten Stufe der DIKW-Pyramide (siehe Abschnitt 2.2). Es basiert auf einer Matrix von zwölf Bewertungsparametern, die in Kombination mit Schulnoten und einer Wichtung eine objektive Bewertung jedes Projektes in

Form einer Gesamtnote ermöglichen. Die Abstufungen der Schulnoten sind ausschließlich ganzzahlig zwischen 1 (sehr gut) und 5 (ungenügend), um klare und vergleichbare Bewertungen zu fördern. Es gibt jeweils Checklisten, was erfüllt sein muss, um die Note 1, 2, 3 oder 4 zu vergeben. Eine 5 gibt es bei Nichterfüllung der Anforderungen für Note 4. Für die Wichtungen siehe Abschnitt 3.2. Die Parameter sind in vier Kategorien gegliedert, die nachfolgend genauer erläutert werden.

3.1 Bewertungsparameter

Die zwölf Bewertungsparameter sind dargestellt in Abbildung 2. Sie bauen häufig auf den Handreichungen und Leitfäden der Arbeitsgemeinschaft BIM4INFRA2020 [6] und den anschließend entwickelten Richtlinien und Musterdokumenten der DB und anderer Behörden auf.

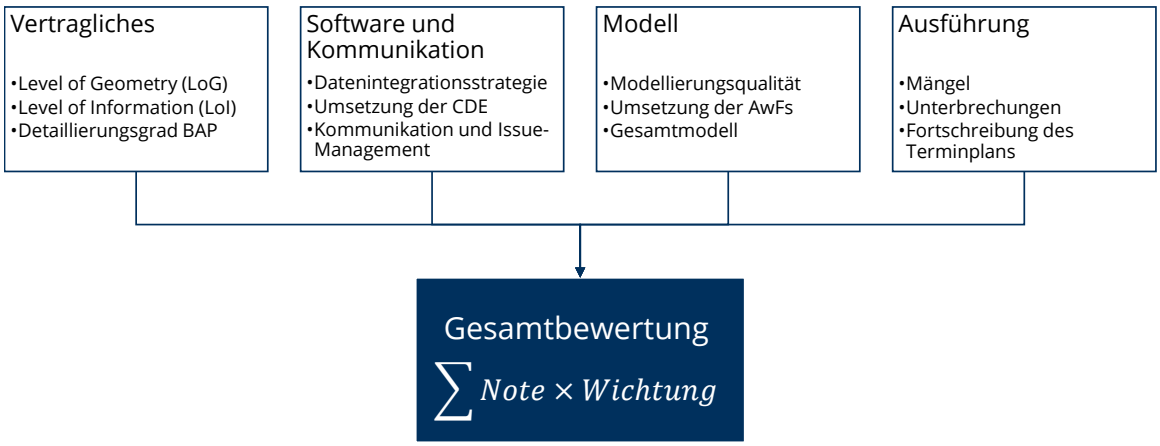


Abbildung 2: Struktur der Bewertungsparameter

3.1.1 Kategorie *Vertragliches*

In den Ausschreibungs- und Vertragsunterlagen gibt es viele Dokumente, die einen erheblichen Einfluss auf das jeweilige Bauprojekt und seine Planungs- und Bauphasen haben. In den beiden Dokumenten AIA und BAP werden die Grundlagen dafür gelegt, dass die BIM-Methodik sinnvoll und reibungslos angewendet wird. Deshalb werden diese auch in den Bewertungsparametern betrachtet und bewertet.

Je weiter ein Projekt voranschreitet, desto höher sollten in den AIA die jeweiligen Anforderungen an die (Fach-)Modelle sein. Nur so kann sichergestellt werden, dass alle für die Ausführung oder den späteren Betrieb relevanten Details modelliert und somit auch bedacht sind. Deshalb sind der geometrische Detaillierungsgrad (*Level of Geometry* - **LoG**) und der Informationsreifegrad (*Level of Information* – **LoI**) jeweils Bewertungsparameter, für die ein Infrastrukturprojekt nur die Note 1 erhält, wenn für jedes Bauteil und für jede Leistungsphase ein angemessener LoG bzw. LoI angefordert ist. Ist der LoG bzw. LoI dagegen nicht für jedes Bauteil und nur für eine Leistungsphase definiert worden, gibt es die Note 3. Wurde überhaupt kein LoG oder LoI vorgegeben, dann wird die Note 5 vergeben.

In den AIA werden weitere Aspekte der BIM-Nutzung vorgeschrieben, diese sind allerdings im BAP detaillierter beschrieben. Deshalb fasst der dritte Parameter aus der Kategorie *Vertragliches* diese Aspekte zusammen: der Parameter **Detaillierungsgrad des BAP** beantwortet folgende Fragen: „Wurden alle Inhalte eines BAP projektspezifisch und ausreichend detailliert beschrieben? Wurde der BAP fortgeschrieben?“ Nur wenn der BAP projekt- und ausschreibungsspezifisch ist, bei signifikan-

ten Änderungen im Projekt fortgeschrieben wurde, die BIM-Anwendungsfälle (AwFs) gelistet und genauer beschrieben sind, digitale Grundlagen, Liefergegenstände und die Lieferzeitpunkte definiert wurden, Rollen und Kommunikation klar erläutert sind sowie Technologien, Modellstruktur und -inhalte beschrieben wurden, gibt es die Note 1. Diese Anforderungen orientieren sich am Muster-BAP der Arbeitsgruppe BIM4INFRA2020 [6] und an den Muster-BAPs der DB InfraGO [7], [8]. Werden nur Teile der Anforderungen oder diese nur unzureichend erfüllt, gibt es eine Abstufung der Note.

3.1.2 Kategorie *Software und Kommunikation*

Die weiteren Bewertungsparameter haben einen ähnlichen Aufbau und ähnliche Noten-Checklisten wie die drei Parameter aus der Kategorie *Vertragliches*. So wurden für die Kategorie *Software und Kommunikation* die Fragestellungen nach Abbildung 3 definiert.

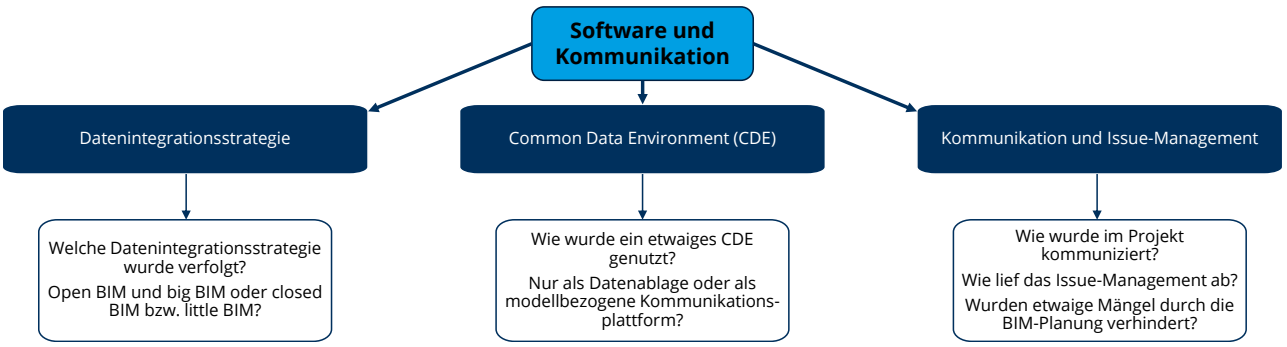


Abbildung 3: Fragestellungen für Kategorie Software und Kommunikation

3.1.3 Kategorie *Modell*

Die Qualität der Entwurfs- und Ausführungsplanung beeinflusst maßgeblich die Ausführung. Nur, wenn die Modellierungsqualität im BIM-Modell angemessen hoch ist, die BIM-Anwendungsfälle (AwFs) umgesetzt wurden und alles in ein Gesamtmodell zusammengeführt wurde lassen sich die Planungen auch fehlerfrei ausführen. Für die Fragestellungen aus der Kategorie *Modell* siehe Abbildung 4.

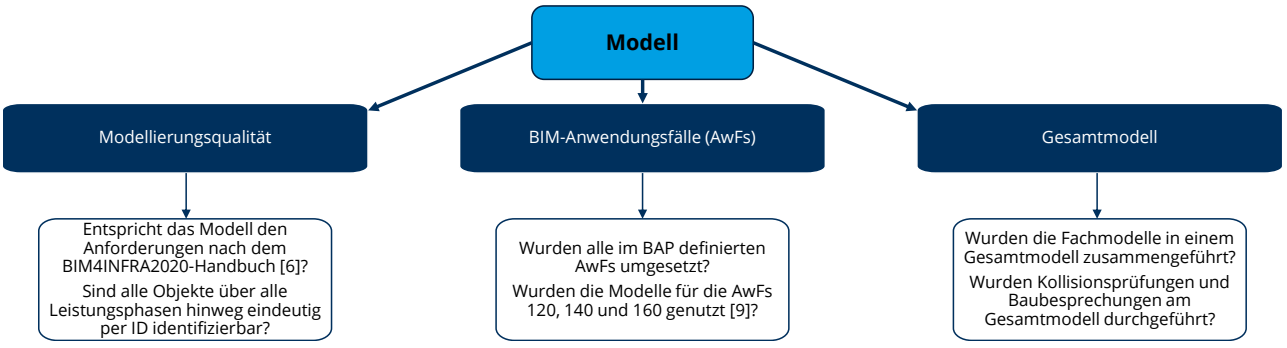


Abbildung 4: Fragestellungen für Kategorie Modell [6], [9]

3.1.4 Kategorie *Ausführung*

Die Kategorie *Ausführung* soll bewerten, welchen Einfluss die zuvor bewerteten Aspekte der Ausschreibung, Planung und Kommunikation auf die tatsächliche Bauausführung hatten. Die Parameter und Fragestellungen sind in Abbildung 5 dargestellt.

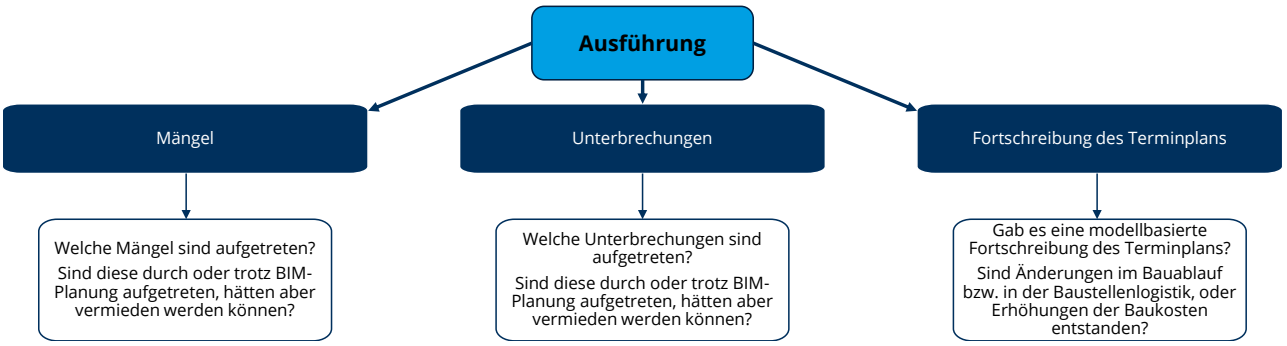


Abbildung 5: Fragestellungen für Kategorie Ausführung

3.2 Wichtungen

Für eine vergleichbare Analyse verschiedener Infrastrukturprojekte mithilfe des zuvor beschriebenen Datenanalysekonzepts wurden Schulnoten eingeführt. Diese können mittels der Wichtungen aus Tabelle 1 zu Kategorienoten und zu einer Gesamtnote für das Projekt verrechnet werden. Die Wichtungen wurden in Abstimmung mit BIM-Experten aus dem Projektteam und der DB InfraGO AG gefunden.

$$\text{Gesamtnote} = \sum \left(\sum \text{Note Parameter} \times \text{Wichtung Parameter} \right) \times \text{Wichtung Kategorie}. \quad (1)$$

Tabelle 1: Wichtungen der Bewertungsparameter

Kategorie	Wichtung Kategorie	Bewertungsparameter	Wichtung Parameter	Wichtung gesamt
Vertragliches	10 %	LoG	25 %	2,5 %
		LoI	25 %	2,5 %
		BAP Detaillierungsgrad	50 %	5,0 %
Software und Kommunikation	35 %	Datenintegrationsstrategie	20 %	7,0 %
		Umsetzung der CDE	40 %	14,0 %
		Kommunikation und Issue-Management	40 %	14,0 %
Modell	30 %	Modellierungsqualität	50 %	15,0 %
		Umsetzung der AwFs	40 %	12,0 %
		Gesamtmodell	10 %	3,0 %
Ausführung	25 %	Mängel	50 %	12,5 %
		Unterbrechungen	25 %	6,3 %
		Fortschreibung des Terminplans	25 %	6,3 %
SUMME	100 %			100 %

4 Fazit

Die Anwendung von BIM in der Bauausführung von Infrastrukturprojekten besitzt große Potenziale, die Ausführung zu optimieren und zu beschleunigen. Dafür muss die BIM-Anwendung in diesem Bereich allerdings standardisiert werden. Für eine erfolgreiche Standardisierung ist eine Analyse der aktuellen Problemstellen und Ableitung von Optimierungspotenzialen notwendig. Für diese Analyse wurde ein Datenanalysekonzept entwickelt, mithilfe dessen BIM-Pilotprojekte aus dem Schieneninfrastrukturbau evaluiert werden sollen. Das Datenanalysekonzept basiert auf zwölf qualitativen Bewertungsparametern, die nach abgeschlossener Analyse eines Infrastrukturprojekts jeweils mit einer Schulnote

zwischen 1 (sehr gut) und 5 (ungenügend) bewertet werden. Durch Checklisten für die Noten wird eine objektive Bewertung und eine einfache Vergleichbarkeit zwischen verschiedenen Projekten garantiert. Mithilfe von Wichtungen lassen sich auch Noten für die Parameterkategorien und das Gesamtprojekt finden. Damit lässt sich leicht überblicken, in welchen Bereichen die BIM-Anwendung bereits gut läuft und wo bei dem jeweiligen Projekt noch Nachholbedarf besteht.

Zu Beginn des Projektes war das Ziel die Konkretisierung der Bewertungsparameter zu Key Performance Indicators (KPIs), also Kennzahlen zur Messung der Effizienz der BIM-Anwendung. Allerdings lässt sich keiner der gefundenen Bewertungsparameter rein mit Kennzahlen bewerten. Es ist immer eine Beschreibung oder eine Auswahl aus einer Liste notwendig. Dadurch wird die Nutzung als KPIs verhindert. Die Bewertungsparameter können auch nicht alle Aspekte beschreiben, die Einfluss auf die BIM-Planung und -Ausführung haben. Dafür ist der Planungs- und Bauprozess zu komplex. Die Parameter sind allerdings eine Ansammlung an Fragestellungen, die die wichtigsten Einflussfaktoren in diesem Bereich behandeln. Dabei wurde sich auf den Einfluss der BIM-Methodik auf Kosten, Bauzeiten und die Entstehung von Mängeln konzentriert. Es wäre sinnvoll, weitere Aspekte und damit auch zusätzliche Parameter in die Bewertung zu integrieren. Dies ist aufgrund des Arbeitsaufwandes für die Analyse und der begrenzten Laufzeit des Forschungsprojekts „BIM2Build“ nicht möglich gewesen, ist aber für eine potentielle Erweiterung des Projekts angedacht. Die hier dargelegten Parameter bilden allerdings die Grundlage für eine erfolgreiche BIM-Anwendung in Infrastrukturprojekten. Die Bewertungsparameter wurden mit Verantwortlichen der DB InfraGO abgestimmt.

Unter den zu analysierenden Projekten befinden sich auch laufende Projekte. Bei diesen kann die Kategorie *Ausführung* nicht im selben Umfang wie bei abgeschlossenen Projekten bewertet werden. Daher werden die Gesamtnoten nicht für den Vergleich herangezogen. Es wird nur mithilfe der Kategorienoten und Parameterwertungen verglichen.

5 Ausblick

Alle Bewertungsparameter lassen sich auf alle Arten von Infrastrukturprojekten anwenden. Nicht nur das, sie sind so gewählt, dass sie sogar im Hochbau angewandt werden können. In diesem Fall müssen allerdings die Dokumente aus der Kategorie *Vertragliches* mit anderen Musterdokumenten und Leitfäden verglichen werden. Auch die AwFs unterscheiden sich zwischen verschiedenen Infrastrukturbetreibern, aber auch zum Hochbau. Ebenso gibt es deutliche Unterschiede zwischen den Modellierungsvorgaben des Parameters „Modellierungsqualität“ für den Hochbau und den Vorgaben der Arbeitsgruppe BIM4INFRA2020 [6]. Grundsätzlich sind die hier beschriebenen Kategorien aber für alle Bauprojekte relevant, in denen BIM zum Einsatz kommt.

In zukünftigen Schritten wird jedes der Schieneninfrastrukturprojekte mithilfe des Konzepts analysiert und bewertet. Anschließend wird mithilfe der vergebenen Noten und Parameter ein Vergleich der Projekte durchgeführt, in dem ermittelt wird, in welchen Bereichen noch Potenziale zur Optimierung der BIM-Anwendung in der Bauausführung bestehen. Diese Optimierungspotenziale werden abschließend zu Handlungsempfehlungen und zu Gestaltungshinweisen für BIM-Modelle und Informationsaustauschprozesse umgewandelt. Damit soll die Standardisierung von BIM in der Bauausführung von Infrastrukturprojekten und somit eine Digitalisierung der Bauprozesse vorangetrieben werden.

Danksagung

Das Forschungsprojekt „Wissenschaftliche Begleitung der BIM-Prozesse des Baus aktueller Schienen-Infrastrukturprojekte – BIM2Build" wird vom Bundesministerium für Digitales und Verkehr im Rahmen des Förderprogramms mFUND (FKZ: 19F1218A) gefördert.

Literatur

[1] Bundesministerium für Verkehr und digitale Infrastruktur, Hrsg., *Stufenplan Digitales Planen und Bauen*, de, Dez. 2015. Adresse: https://bmdv.bund.de/SharedDocs/DE/Publikationen/DG/stufenplan-digitales-bauen.pdf

[2] DB AG, »BIM-Strategie Implementierung von Building Information Modeling (BIM) im Vorstandsressort Infrastruktur der Deutschen Bahn AG«, de, Feb. 2022. Adresse: https://www.deutschebahn.com/resource/blob/6876006/c4ae3a7c344b770a0e762aa73406598c/Implementierung-von-BIM-im-VR-I-data.pdf

[3] DB InfraGO AG, *BIM2Build - Das Testprogramm für BIM in der Bauausführung*, de. Adresse: https://bim.dbinfrago.com/bim-fahrweg/BIM-Weiterentwicklung/BIM2Build-Das-Testprogramm-fuer-BIM-in-der-Bauausfuehrung-13023790

[4] R. L. Ackoff, »From data to wisdom«, en, *Journal of applied systems analysis*, 1989.

[5] J. Rowley, »The wisdom hierarchy: representations of the DIKW hierarchy«, en, *Journal of Information Science*, Jg. 33, Nr. 2, S. 163–180, Apr. 2007. DOI: 10.1177/0165551506070706

[6] Arbeitsgruppe BIM4INFRA2020, *Handreichungen BIM4INFRA2020*. Adresse: https://www.bim4infra.de/

[7] DB InfraGO AG, Geschäftsbereich Fahrweg, *Musterdokument BAP - BIM-Projektabwicklungsplan*, de, Dez. 2023. Adresse: https://www1.deutschebahn.com/resource/blob/12624362/8fe92b06fe672df3a1e5c3b4a44c1414/DB-InfraGO-Fahrweg-BAP-V-6-0-extern-data.pdf

[8] DB InfraGO AG, Geschäftsbereich Personenbahnhöfe, *BIM–Projektabwicklungsplan (BAP) - Vorlage*, de, Feb. 2025. Adresse: https://infoplattform-personenbahnhoefe.deutschebahn.com/resource/blob/7716938/c2f9cc44581ed217c8dcdec975bdec5f/BIM-Projektabwicklungsplan-data.pdf

[9] DB InfraGO AG, *BIM-Anwendungsfälle*, de. Adresse: https://bim.dbinfrago.com/bim-fahrweg/BIM-Standards/BIM-Anwendungsfaelle-5536960

Virtual Masonry Shear Wall Experiments: Enhancing Understanding of Nonlinear Behavior through Virtual Simulation

Raymond Leonardo Chandra[1], Stephan Dreyer[2], Florian Kolisch[2], Koen Castermans[1], Djamel Berkaoui[1], Patrick Querl[1], Heribert Nacken[1] and Sven Klinkel[2]

[1]Academic and Research Department Engineering Hydrology and UNESCO Chair of Hydrological Change and Water Resources Management (LFI), RWTH Aachen University, Aachen, Germany

[2] Institute of Structural Analysis and Dynamics (LBB), RWTH Aachen University, Aachen, Germany

E-mail(s): chandra@lfi.rwth-aachen.de, dreyer@lbb.rwth-aachen.de, kolisch@lbb.rwth-aachen.de, castermans@lfi.rwth-aachen.de, berkaoui@lfi.rwth-aachen.de, querl@lfi.rwth-aachen.de, nacken@lfi.rwth-aachen.de, klinkel@lbb.rwth-aachen.de

Abstract: In nonlinear structural analysis, civil engineering students are required to understand the structural performance of shear walls under various loading conditions, particularly lateral forces such as those caused by wind or seismic activity. Physical shear wall testing helps engineers to understand damage and failure, which is essential for safe structural design. However, building a shear wall for testing demands time and resources, as the tests are inherently destructive. Additionally, students often lack access to the specialized equipment needed for shear tests, and safety precautions prevent them from conducting experiments unsupervised. As a result, they typically rely on passive learning methods such as lectures or textbooks. In this project, Virtual Reality (VR) Masonry Lab, we created an open-source multiplayer application that allows students to construct and test masonry shear walls in a virtual environment. In the application, students can collaboratively design masonry walls by specifying dimensions and brick size. They then apply horizontal loads to simulate a pushover test, gradually increasing force until the wall fails. The algorithm calculates the wall's deformation, damage, and failure based on the input parameters, and the results are visually rendered in real-time, which allows students to observe how design parameters can influence the structure. This VR-based approach promotes trial-and-error experimentation, allowing students to explore how design parameters affect wall behavior under stress. The digital environment eliminates manual construction and material waste, enabling safe, repeatable testing without time constraints. In a pilot classroom session, students reported that the tool improved their understanding of nonlinear structural behavior and appreciated the ability to virtually repeat the experiment from anywhere and anytime.

Keywords: Virtual Reality, Masonry Shear Walls, Structural Education, Nonlinear Behavior

DOI: 10.18154/RWTH-CONV-254908. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

The structural performance of masonry shear walls under lateral loading is a critical topic in earthquake-resistant design [1]. Understanding how these walls respond to seismic forces, including deformation and failure mechanisms, is essential for designing safe and resilient buildings [2]. However, traditional experimental methods used to investigate this behavior are time-consuming, resource-intensive, and often inaccessible to students due to safety and equipment constraints [3].

In civil engineering education, especially in nonlinear structural analysis, students often rely on theoretical knowledge and passive learning tools such as textbooks and lectures [4]. While these approaches convey foundational concepts, they are often insufficient to illustrate complex nonlinear behavior of masonry structures under real loading conditions [5]. This gap limits students' understanding of how design parameters affect structural performance [5].

To address this issue, an open-source Virtual Reality (VR) application called the VR Masonry Lab has been developed. In this system, structural calculations are performed by the EasyFEM finite element framework, while Unity handles the immersive visualization of deformation, crack propagation, and collapse based on user-defined parameters. This allows students to interactively construct walls, apply horizontal loads, and observe the resulting nonlinear behavior in a simulated pushover test.

The VR Masonry Lab builds on the MyScore project, which promotes avatar-based teaching in higher education. MyScore offers an open-source framework that enables interaction via avatars in immersive 3D scenarios [6]. Its modular design, support for real-time feedback and collaborative learning via multiplayer made it an ideal foundation for building the VR Masonry Lab.

Integrating VR into structural education bridges the gap between theory and practice. It enables hands-on exploration in a risk-free digital environment, supporting deeper insight into nonlinear behavior and failure modes. The simulation uses validated numerical models from the EasyFEM to ensure accuracy and relevance. This paper presents the methodology, simulation framework, and educational application of the VR Masonry Lab.

2 Related Work

The use of computer technologies in structural engineering education and research has grown significantly [7], [8], [9]. Among them, VR has shown promise in helping students understand complex structural behavior by interacting with digital environments that simulate real-world responses [7], [9]. Several studies have explored VR in civil engineering education. One study presented a VR beam analysis module where users applied loads and received real-time visual feedback on deformation and stress distribution [9]. Results showed improved learning outcomes compared to traditional lecture [9]. VR labs have been explored as alternatives to physical labs in engineering education [10]. While students valued the accessibility and clarity of VR, challenges remained with hands-on skill transfer [10]. VR has been used in structural design to support parametric modeling of concrete shear walls, enabling visualization of complex geometries and reinforcement details often difficult to handle with conventional tools [8]. Other work has applied VR to assess damage in heritage structures and visualize seismic degradation, aiding both preservation and structural education [7]. Augmented reality

(AR) has also addressed practical issues in masonry construction, such as embed coordination and material planning which help in reducing errors and waste before execution [11].

While prior efforts focus on visualization and design, few platforms allow users to build, load, and observe structural failure interactively. The VR Masonry Lab fills this gap by combining finite element-based simulation with immersive interaction. This application allows students to conduct virtual pushover tests and explore nonlinear behavior under varying design parameters in a safe, repeatable, and collaborative environment.

3 Methodology

The VR Masonry Lab was developed as an open-source application aimed at simulating the nonlinear behavior of masonry shear walls under horizontal loading. The methodology integrates structural simulation, virtual interaction, and educational usability within a single framework. This section outlines the simulation model, virtual interface design, and the workflow for user interaction and analysis.

3.1 Physical Masonry Lab

Masonry shear walls are structural elements commonly used in buildings to resist horizontal loads caused by wind or seismic events [12]. They are typically constructed from bricks or concrete blocks arranged in a regular pattern and bonded together with mortar [12]. These walls provide lateral stiffness and stability of a structure through in-plane shear resistance [2]. In unreinforced masonry, the behavior is highly nonlinear and often brittle, with failure modes that include diagonal tension cracking, corner crushing, and sliding along mortar joints [2]. When properly designed, masonry shear walls can efficiently dissipate energy and limit lateral displacements during seismic activity [12].

In a laboratory setting, masonry shear walls are tested to evaluate their mechanical behavior under horizontal loading conditions, particularly to simulate seismic effects [13]. Figure 1 illustrates a typical experimental setup used for such testing.

The wall specimen is first constructed according to standard geometries and material specifications [2], then mounted onto a strong laboratory floor or reaction frame with rigid boundary conditions at the base. A hydraulic actuator applies lateral forces at the top of the wall and is connected to a steel reaction frame that resists the applied load. In many cases, vertical confinement is also introduced using steel beams to simulate axial forces from upper building stories [12]. Additional instrumentation, such as displacement sensors, load cells, and strain gauges, is installed to monitor the wall's response during testing.

The lateral load is incrementally increased in what is referred to as a pushover test [12] which allows researchers to observe progressive cracking, deformation, and ultimately structural failure. This controlled environment enables the detailed study of load-displacement behavior, crack development, and energy dissipation characteristics, all of which are critical for understanding the seismic performance of masonry walls. To illustrate such failure, figure 2 shows a masonry shear wall after undergoing a lateral loading. Typical shear failure patterns such as diagonal and stepped cracks are visible and have been marked in the image for clarity. This crack pattern is characteristic of in-plane seismic loading, where the wall gradually deforms until failure.

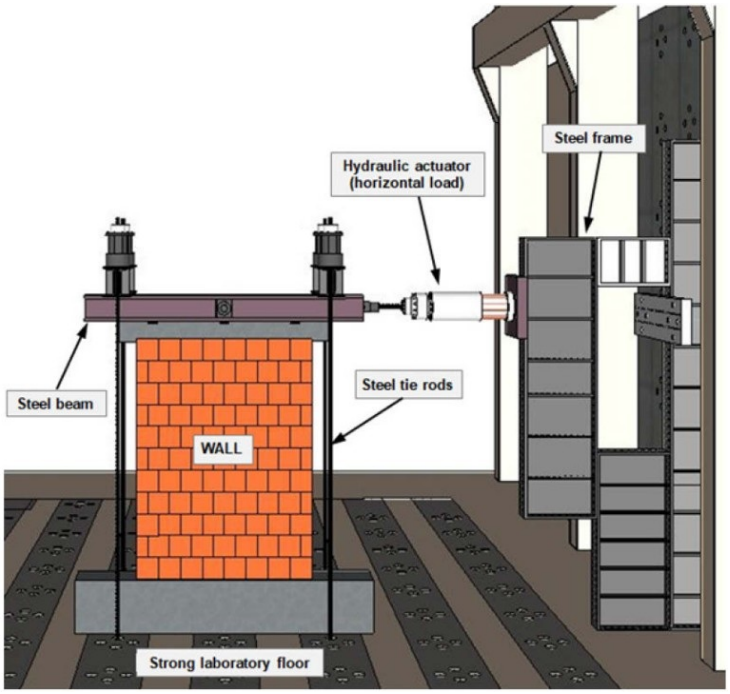


Figure 1: Wall test schematic [12]

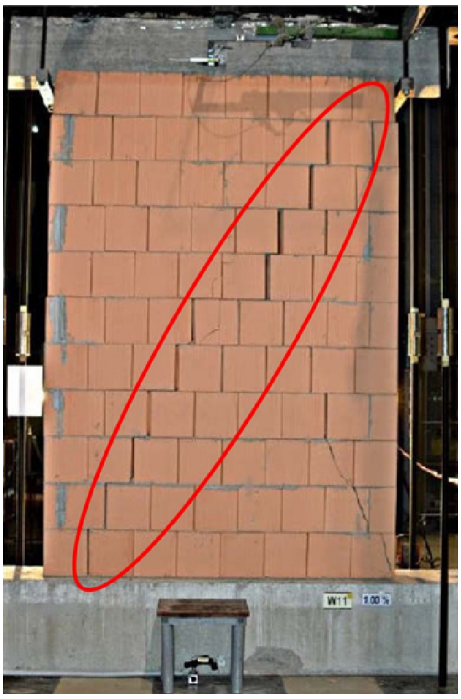


Figure 2: Failed masonry wall [12]

While these tests offer valuable insights, they demand significant equipment, materials, time, and safety precautions. Their destructive nature also limits their suitability for educational settings, where repeated, hands-on experimentation is often preferred [9].

3.2 Virtual Masonry Lab Overview

To address the limitations of physical testing and improve accessibility, the VR Masonry Lab was designed to replicate masonry shear wall experiments within an interactive virtual environment. The application was built using the Unity game engine and is based on the open-source MyScore framework, which supports multi-user, avatar-based learning in immersive 3D spaces. The structural response of the masonry wall was computed using EasyFEM, a finite element library for structural analysis developed by the Chair of Structural Analysis and Dynamics (LBB). Simulation data from EasyFEM is passed to Unity, which visualizes the wall's deformation, cracking, and collapse behavior in real time. This setup ensures accurate structural calculations with an intuitive user interface. The application runs on VR headsets and Windows. Figure 3 shows the virtual lab setup, where two avatars (representing users) are present in the scene. The environment features a masonry shear wall testing rig at the center, along with a control panel for user input and a real-time force-displacement graph. A 360° video of a physical laboratory is embedded in the virtual space to visually connect real and digital environments.

The experiment begins with users specifying brick and wall properties via the control panel, including brick dimensions (height, width, mass) and wall geometry (height and length) as shown in figures 4 and 5. Selected values are summarized on the left panel for review before starting the simulation.

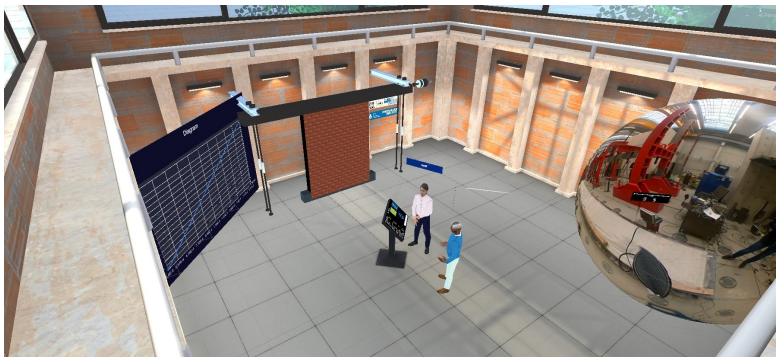


Figure 3: Virtual Masonry laboratory

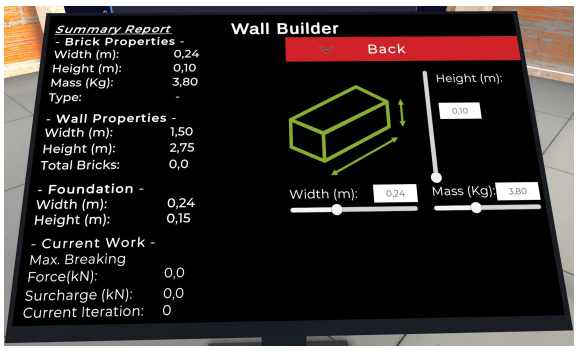


Figure 4: Brick configuration

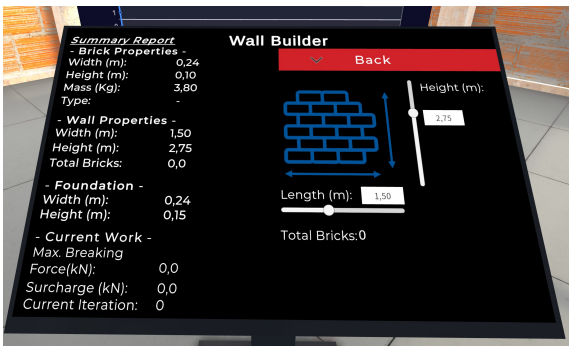


Figure 5: Wall configuration

Currently, the system supports unreinforced clay bricks bonded with mortar. The bricks have a density of approximately 1800 kg/m^3 and an elastic modulus of 3000–5000 MPa. The mortar provides a tensile bond strength between 0.1–0.4 MPa.

Users can then initiate the simulation by selecting the "Build" function as shown in figure 6. The system automatically assembles the wall brick by brick based on the defined specifications. The wall construction animation serves only as a visual feature and does not affect structural calculations.

Figure 6: Constructed virtual wall

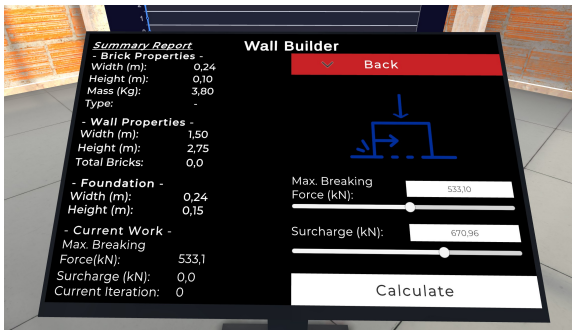


Figure 7: Acting force menu configuration

Afterwards, users can define horizontal loading parameters such as breaking force (kN) and surcharge (kN) via the "Acting Work" menu (figure 7). When the "calculate" function is selected, the configured force will be applied to the top of the wall to perform a virtual pushover test. The EasyFEM engine applies the defined loading to compute internal stresses, displacements, and failure progression.

Next, the simulation engine visualizes wall response through displacement, cracking (figure 8), and eventually, full collapse when the applied force exceeds the wall's load-bearing capacity (figure 9). The simulation stops once the wall exceeds its load-bearing capacity, with broken bricks scattered to indicate failure.

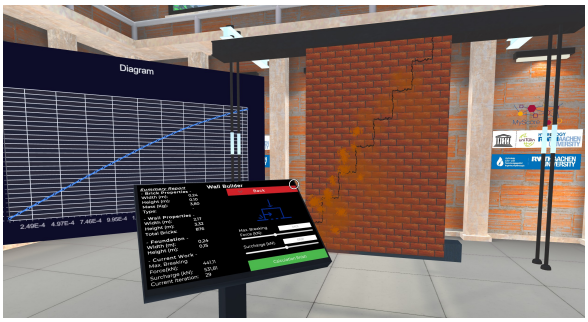


Figure 8: Wall cracking

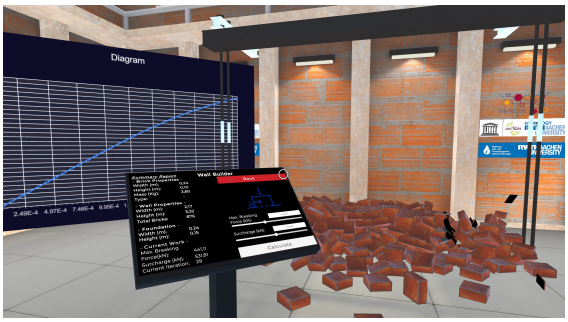


Figure 9: Collapsed wall

A live force-displacement diagram adjacent to the wall illustrates the structural response during loading. Initially, the curve rises steadily, indicating elastic deformation. As loading increases, cracks appear, the curve peaks, and failure is both numerically and visually represented.

To encourage exploration and experimentation, users can repeat the process with different wall configurations. This interactive workflow enables students to test hypotheses and observe how small changes in design or material properties can lead to significant differences in structural behavior. Through its immersive design, the VR Masonry Lab deepens understanding of nonlinear behavior and structural design decisions.

3.3 Student Test

To evaluate the educational potential of the VR Masonry Lab, the application was demonstrated during a lecture session of the "Nonlinear Structural Analysis" course at RWTH Aachen University, with approximately 50 civil engineering students in attendance. The goal was to show how the simulation supports understanding of shear wall behavior and design sensitivity. Although only a few students interacted directly with the system, the entire class followed the simulation via live screen casting from the VR headset to a monitor, allowing everyone to observe the setup, loading process, and wall response. The pilot was conducted in a single-user mode due to classroom logistics. Regardless, the VR Masonry Lab supports true multiplayer functionality, enabling users to collaborate remotely in the same virtual environment.

Students were briefly introduced to VR controls before experimenting with different wall and brick configurations then applied lateral loads and observed force-displacement responses and cracking. Feedback collected at the end of the session was highly positive. Students found the application more intuitive than traditional diagrams and appreciated the ability to repeat experiments without time constraints. Many also valued the flexibility of using the tool on desktop devices and its support for remote collaboration.

Future evaluations will include a structured learning outcome study, comparing groups using traditional methods with those using the VR application. Planned methods include pre- and post-testing, usability

surveys, and focus group interviews. These steps aim to assess not only user experience but also the tool's impact on conceptual understanding. Furthermore, students suggested several improvements to the application, including graphical improvement and the ability to define additional parameters such as vertical loads.

Overall, the pilot session demonstrated the advantages of VR simulation in engineering education. By allowing students to build and test structures in a 3D environment, the VR Masonry Lab bridges the gap between theory and practice, supports intuitive understanding of nonlinear behavior, and offers an exploratory learning experience.

4 Conclusions and Future Work

This paper addressed a key challenge in structural education: the limited accessibility of hands-on shear wall testing due to safety, resource, and equipment constraints. As shown in prior research, existing VR tools in engineering education often focus on design visualization but lack real-time simulation or interactive experimentation. The VR Masonry Lab bridges this gap by providing a virtual environment where students can configure, load, and observe the behavior and failure of masonry shear walls.

By simulating a pushover test in an immersive environment, the software enables students to explore nonlinear structural behavior in a repeatable, risk-free format. Unlike purely visual VR tools, the VR Masonry Lab includes real-time structural feedback, enabling a deeper exploration of failure behavior. The pilot classroom test confirmed that students found the tool engaging and educational, offering a more intuitive understanding than traditional methods. The ability to collaborate remotely and experiment with various configurations added further value.

Future work will focus on expanding the simulation capabilities, including the addition of vertical and dynamic loading, additional material types, and refined failure modeling. In parallel, a structured user study will be conducted to evaluate learning outcomes and usability. The study will involve controlled comparison between students using the VR Masonry Lab and those following conventional teaching methods. Assessment will include pre- and post-tests, user satisfaction surveys, and observation of in-VR behavior to measure engagement and conceptual understanding. These results will guide further improvements and validate the educational value of the tool.

5 Acknowledgement

We gratefully acknowledge the use of the RWTHDynLab at RWTH Aachen University, where the masonry shear wall experiments were conducted. The lab was funded by the German Research Foundation (DFG) under grant numbers INST 222/1161-1 FUGG and INST 222/1403-1 FUGG. This project was internally funded by RWTH Aachen University under the ICON – International Cooperation Online initiative.

6 Conflict of Interest

The authors declare that this research was carried out without any commercial or financial interests that could be perceived as a potential conflict of interest.

References

[1] Z. Zhang and F. Wang, “Experimental investigation into the seismic performance of prefabricated reinforced masonry shear walls with vertical joint connections”, en, *Appl. Sci. (Basel)*, vol. 11, no. 10, p. 4421, May 2021.

[2] A. Milijaš, M. Marinković, C. Butenweg, and S. Klinkel, “Experimental investigation on the seismic performance of reinforced concrete frames with decoupled masonry infills: Considering in-plane and out-of-plane load interaction effects”, en, *Bull. Earthquake Eng.*, vol. 22, no. 15, pp. 7489–7546, Dec. 2024.

[3] Z. Wu and X. Chen, “Consumption estimating model of the frame shear wall structure materials”, in *LISS 2014*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 1163–1168.

[4] J. Mills, “A case study of project based learning in structural engineering”, in *2002 Annual Conference Proceedings*, Montreal, Canada: ASEE Conferences, 2020.

[5] I. May and D. Johnson, “The teaching of structural analysis”, *Structural Engineer*, vol. 86, Nov. 2008.

[6] D. Berkaoui, R. Chandra, and K. Castermans, *MyScore – avatar-based teaching and learning*, 2022.

[7] F. Muñoz, F. Peña, and M. Meza, “Virtual reality models for the structural assessment of architectural heritage buildings”, *Int. J. Archit. Heritage: Conserv. Anal. Restor.*, vol. 8, no. 6, pp. 783–794, Nov. 2014.

[8] F. Zhou and J. Sun, “Application of virtual reality technology in parametric design of fabricated shear wall structure”, en, in *Novel Technology and Whole-Process Management in Prefabricated Building*, Singapore: Springer Nature Singapore, 2024, pp. 175–182.

[9] L. Zhang, C. Moen, H. Blum, and B. Marks, *Structural analysis in virtual reality for education with bmly*, https://jscholarship.library.jhu.edu/server/api/core/bitstreams/3174efb8-9491-46b5-aafb-d62b13f5f71e/content, 2022.

[10] R. L. Chandra, D. Berkaoui, P. Querl, K. Castermans, and H. Nacken, “Student perceptions of VR labs modeled after traditional physical labs in non-destructive measurement testing”, *Int. J. Inf. Educ. Technol.*, vol. 15, no. 3, pp. 419–427, 2025.

[11] D. Olsen, J. Kim, and J. M. Taylor, “Using augmented reality for masonry and concrete embed coordination”, in *Proceedings of the Creative Construction Conference 2019*, Budapest University of Technology and Economics, 2019.

[12] N. J. Boesen, *Trag- und Verformungsverhalten von unbewehrten Mauerwerksscheiben unter Berücksichtigung der Interaktion mit der Gebäudestruktur*. RWTH Aachen University, 2021.

[13] A. Milijaš, M. Marinković, C. Butenweg, and S. Klinkel, “Experimental results of reinforced concrete frames with masonry infills with and without openings under combined quasi-static in-plane and out-of-plane seismic loading”, en, *Bull. Earthquake Eng.*, vol. 21, no. 7, pp. 3537–3579, May 2023.

Virtual Structural Analysis of a Shaking Table for Earthquake Simulations

Raymond Leonardo Chandra[1], Stephan Dreyer[2], Florian Kolisch[2], Koen Castermans[1], Djamel Berkaoui[1], Patrick Querl[1], Heribert Nacken[1] and Sven Klinkel[2]

[1]Academic and Research Department Engineering Hydrology and UNESCO Chair of Hydrological Change and Water Resources Management (LFI), RWTH Aachen University, Aachen, Germany

[2] Institute of Structural Analysis and Dynamics (LBB), RWTH Aachen University, Aachen, Germany

E-mail(s): chandra@lfi.rwth-aachen.de, dreyer@lbb.rwth-aachen.de, kolisch@lbb.rwth-aachen.de, castermans@lfi.rwth-aachen.de, berkaoui@lfi.rwth-aachen.de, querl@lfi.rwth-aachen.de, nacken@lfi.rwth-aachen.de, klinkel@lbb.rwth-aachen.de

Abstract: In modern engineering education, bridging the gap between theoretical knowledge and practical application is an important challenge. However, students often have limited access or no access to hands-on experiences, particularly with specialized equipment like shaking tables (devices that simulate how earthquakes can affect building structures), which are not always available in university teaching. Typically, students can only observe such devices during occasional demonstrations by lecturers, through scheduled appointments, or not at all. Consequently, most of the time they rely only on lectures or video material to understand how shaking tables work. In this project, Virtual Structural Analysis, we utilize immersive Virtual Reality (VR) technology to visualize and simulate real-world shaking table experiments. This project provides students with an interactive application in which they can experiment with shaking tables independently. Within the VR environment, students can input various simulation parameters, such as stiffness, frequency, amplitude, etc., to observe how these parameters affect building behavior during earthquakes. The project aims to deepen students' understanding of complex structural dynamics through active experimentation and immersion, as a complement to traditional teaching methods. The pilot testing with 11 students showed that the VR application helped them grasp the different parameters that influence building dynamics. Moreover, the students reported that the program improved their comprehension of the lecture content. A larger study using objective tests and comparisons with traditional instruction will be planned. Future updates will support more structural types and diverse input signals for learning analytics.

Keywords: Shaking Table Experiments, Virtual Reality, Structural Dynamics, Finite Element Analysis

DOI: 10.18154/RWTH-CONV-254879. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Earthquakes are among the most destructive natural hazards and are capable of inflicting widespread damage to buildings and infrastructure [1]. For structural engineers, accurately predicting how buildings

respond to seismic loading is essential to designing safe and resilient structures [2]. Achieving this requires not only theoretical understanding, but also observation of structural behavior under realistic dynamic conditions. [1].

One experimental tool for this simulation purpose is the shaking table. These devices reproduce ground motions by imposing controlled vibrations on scaled structural models, allowing engineers to observe how different systems respond to seismic excitation [1]. Regardless of complexity, from uniaxial to multiaxial, all shaking tables serve to test structural models under simulated earthquake loads. [3].

Despite their value in research, shaking tables are rarely accessible in typical engineering education [4]. Cost, large-footprints, and safety constraints often reduce their use to occasional demonstrations [4], [5]. This often forces students to rely on passive instruction, limiting their ability to visualize real seismic effects. [6].

To bridge this gap, we developed the Virtual Structural Analysis application: a free immersive Virtual Reality (VR) and desktop application that simulates the operation of a shaking table. The simulation replicates a physical 3.0 x 3.0 meter uniaxial seismic simulator from MTS Systems Corporation [7]. Within this virtual environment, users can manipulate stiffness, damping, and mass to see their effects on structural response. By enabling hands-on, parameter-driven experimentation, the system transforms passive observation into active learning [8], [9].

While a desktop version is provided, VR offers several pedagogical advantages. Unlike traditional screen-based tools, VR enables spatial immersion, free exploration of the experimental setup, and multi-angle observation of structural motion. Prior studies suggest that these features can enhance spatial understanding, engagement, and memory retention, particularly for complex three-dimensional phenomena [8], [9].

Our motivation is to offer a practical, engaging, and scalable learning tool that complements classroom instruction. To ensure accessibility, the application is free to download and use. For students without VR hardware, a desktop version is also available for Windows system. The application can be accessed at our *file server*.

Building on this foundation, this paper presents the system's design, educational integration, and initial evaluation.

2 Background and Related Work

Although the use of shaking tables in earthquake engineering is well established and has significantly advanced seismic design, their application in education remains limited due to high costs, space, and safety requirement [2], [5]. As a result, alternative approaches have been explored to improve accessibility, particularly in academic environments.

One direction involves low-cost physical setups, such as open-source uniaxial shaking tables like Namazu and Shakebot, which support programmable control and are compatible with tools like MATLAB or Python. Shakebot, for example, combines motion control with computer vision for real-time feedback, showing promise for classroom use [4], [10]. In addition to hardware solutions, VR systems

simulate earthquake effects on building interiors using 360-degree imagery and pre-recorded seismic data [11]. However, most of these systems are limited to passive experiences, aimed more at public awareness than active engineering education.

In contrast, a VR-based digital twin enabled parameter manipulation and structural visualization in real time [12], while another system visualized nonstructural responses like ceiling failures [13]. These applications highlight the pedagogical potential of immersive interaction. Another study evaluates how VR labs help engineering students learn measurement techniques, finding them useful for concept learning but less effective for hands-on skill development [14].

Compared to these approaches, our project introduces a fully interactive VR shaking table simulation grounded in real experimental setups. Unlike most passive or hardware-limited systems, it supports real-time parameter control, offers both VR and desktop access, and mirrors authentic test procedures, thereby bridging the gap identified in the literature.

3 Material and Methodology

3.1 Physical Shaking Table

The simulation is based on a uniaxial shaking table and structural testing system located at the Chair of Structural Analysis and Dynamics (LBB) at RWTH Aachen University [3]. Developed by MTS Systems Corporation, the physical system, shown in figure 1, served as the reference for digitally replicating both the table's mechanical behavior and the dynamic response of a full-scale test structure [7]. The goal is to enhance accessibility and interactivity in structural dynamics education.

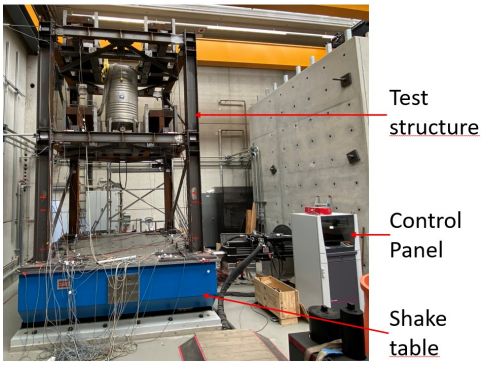


Figure 1: Physical shaking table

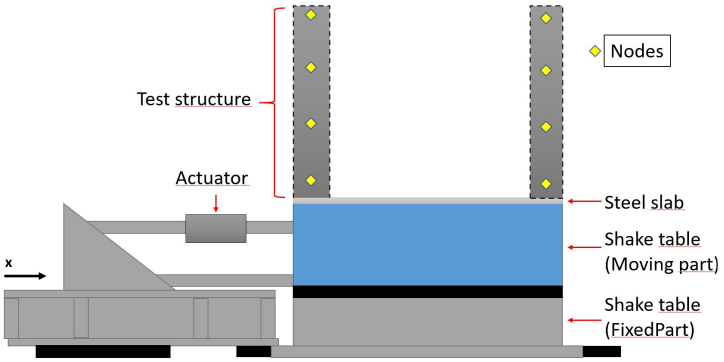


Figure 2: Schematic of a shaking table

The shaking table features a 3.0 × 3.0 meter platform, supports specimens up to 10 metric tons, and operates with a maximum actuator stroke of ±250 mm. It can reach velocities of ±1 m/s, accelerations of ±1g at rated load, and frequencies between 0 and 50 Hz, while withstanding overturning moments up to 30 meter-tons.

Figure 2 illustrates the uniaxial setup used for dynamic testing. A hydraulic actuator drives the platform horizontally to reproduce ground motion. The test specimen is a multi-storey steel frame fixed to the top of the shaking table. The frame columns are instrumented with multiple accelerometers, shown in figure 2 as yellow diamond symbols. These sensors are installed at discrete nodes along the height of the structure. Each node represents a measurement point corresponding to a specific floor level

or elevation within the structural model. The sensors capture acceleration profiles and structural response during seismic loading. The blue platform represents the shaking table itself, which transmits the input motion directly to the base of the specimen. The entire setup is mounted on a stiff foundation to ensure minimal energy loss and accurate reproduction of input dynamics. This setup serves as the reference for the virtual simulation.

3.2 Virtual Shaking Table

The virtual shaking table simulation was developed using the Unity game engine. All 3D models and the virtual environment were created in Blender. The simulation's structural and shaking table responses are computed using EasyFEM, a finite element library developed by LBB. The virtual scenario is placed in a minimalistic laboratory space devoid of unnecessary visual elements. Figure 3 shows the appearance of the virtual simulation, including the digitally reconstructed shaking table and its surrounding interface.

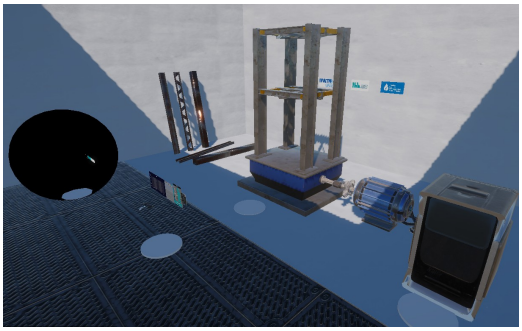


Figure 3: VR shaking table scene overview

As shown in figure 4, users interact with the shaking table via the main control panel, where (in the current version) they can adjust the following parameters:

- Emodul (Elastic Modulus): This input defines the material stiffness (kN/m^2).
- Flächenträgheitsmoment (Moment of Inertia): This parameter influences the resistance of beams and columns to bending (m^4).
- Punktmasse 1 and Punktmasse 2 (Point Mass 1 and 2): These fields define concentrated masses (kg) located at discrete heights within the structure.
- Rayleigh Systemdämpfung 1 and 2 (Rayleigh Damping Coefficients): These inputs correspond to the damping ratios associated with specific vibration modes ($\%$). They are used in Rayleigh damping formulations to simulate energy dissipation mechanisms in the structure.
- Sinusanregung (Sinusoidal Ground Excitation Amplitude): This field sets the amplitude (m) of the sinusoidal ground motion applied by the virtual shaking table.
- Frequenz (Frequency): This parameter determines the frequency of the sinusoidal input (Hz). It allows users to investigate frequency-dependent behavior and resonance phenomena by exciting the structure at various rates.

- Amplitude: This output field displays the calculated response amplitude of the system based on the selected input parameters.

While the current version does not restrict input ranges, future updates will introduce suggested value limits based on typical engineering scenarios, to help users avoid unrealistic configurations and ensure meaningful simulation results.

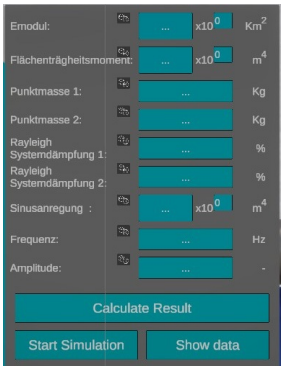


Figure 4: User menu

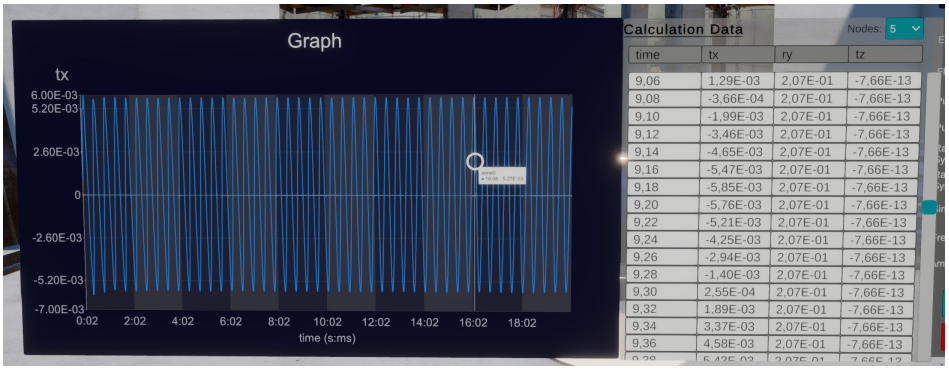


Figure 5: Extended user menu: data and graph visualization

After defining the parameters, users begin the simulation by clicking *Calculate Result*, which computes the displacement of each structural node based on the input values. Pressing *Start Simulation* animates the structure's real-time response on the virtual shaking table as shown in figure 6.

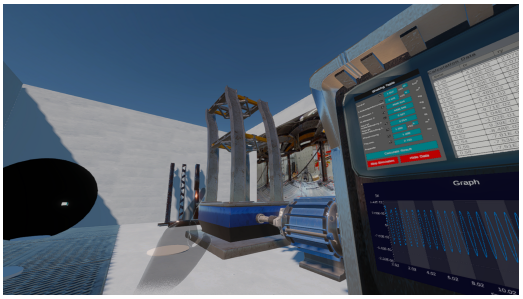


Figure 6: The VR shaking table performs simulations based on user-defined parameters

Users can also choose the *Show Data* to view X, Y, and Z displacement values in a structured table. They can also inspect displacement at individual nodes. Data is sampled every 0.2 seconds for 20 seconds and visualized in table and graph form (figure 5).

To bridge the gap between simulation and reality, a video of the corresponding physical shaking table test is displayed behind the virtual setup (figure 7), allowing direct comparison between simulated and real-world behavior.

Moreover within the VR environment, students have the option to explore a 360° video recording of a real shaking table experiment. The "bubble view," shown in figure 7 allows immersive observation of the real lab.

For accessibility, a desktop version is available for users without a VR headset. VR users may interact using standard controllers or hand tracking for a more natural experience.

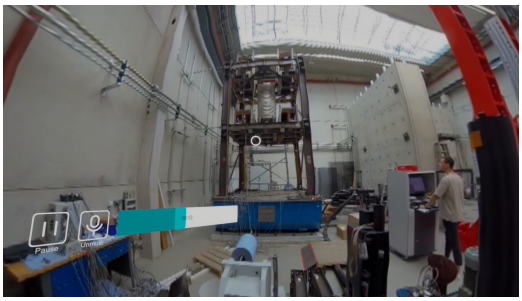


Figure 7: 360° bubble view of a physical shake table

4 Application Testing

In order to evaluate both usability and instructional effectiveness of the VR-based shaking table simulation, a standardized questionnaire was distributed after the user testing session. The survey included seven Likert-scale statements and six open-ended questions. Participants ($n = 11$) responded on a six-point scale (1 = strongly agree, 6 = strongly disagree).

Overall, feedback was highly positive, particularly for learning and ease of system use. The simulation helped students understand how structural parameters affect dynamic behavior ($\mu = 1.64$). Similarly, the ability to understand how vibrations get transferred through a structure was rated positively ($\mu = 1.82$). Participants reported improved understanding of lecture and exercise content after using the simulation ($\mu = 2.0$) which shows its effectiveness as a complementary instructional tool.

The simulation was rated visually realistic ($\mu = 1.25$). Although user experience was positive, about half reported motion sickness. ($\mu = 3.0$). Though teleportation and vignette were included, many users missed or avoided them, suggesting a need for clearer access or better defaults. Notably, most participants stated they would recommend the simulation as part of the lecture ($\mu = 1.5$). These results show that the simulation is informative, engaging, and user-friendly. The findings support integrating immersive technologies into structural engineering courses and provide a strong foundation for future improvements.

In addition to the rating-based feedback, participants answered six open-ended questions:

- *What did you find particularly helpful or informative in the program?* Many students highlighted the freedom to walk around the setup, observe vibrations from various angles, and adjust parameters in real time. This interactivity was frequently cited as essential for deepening conceptual understanding.
- *What did you find unclear or difficult to understand?* Some users noted unclear parameter labels and missing units; a few were unsure about specific interface elements.
- *Which features would you like to see improved?* Common suggestions included movable/closable input windows, tooltips, and clearer navigation instructions.
- *What additional features would you find helpful in future versions?* Participants expressed interest in including multiple structural configurations, support for various excitation types (e.g., impulse or real earthquake data), and functionality to compare different simulations side by side.

- *What was your impression of the 360 ° bubble view?* Most found the 360 ° video view immersive and informative for understanding real-world shaking table experiments. However, a few users felt it slightly affected their spatial orientation in VR.

While these initial results are encouraging, the study has clear limitations. The sample size was small and based on volunteer participation, and the evaluation relied primarily on subjective feedback. To establish learning effectiveness more rigorously, a follow-up study is planned that will involve a larger cohort of students, a control group using traditional media (e.g., lecture slides or videos), and objective measures such as pre- and post-tests or delayed retention quizzes. Additionally, the feedback provides clear direction for future improvements, especially in usability, simulation options, and user guidance.

5 Conclusions and Future Work

This study introduced a Virtual Reality (VR) simulation to replicate shaking table experiments for structural dynamics education. The tool allows students to adjust structural parameters and observe seismic responses interactively. Unlike passive or hardware-limited systems, it offers an immersive, scalable, and accessible platform to bridge theoretical knowledge and practical understanding.

Preliminary testing with 11 participants indicated its potential effectiveness. Quantitative results showed strong agreement on learning impact and usability, while qualitative feedback praised its interactivity, visualization clarity, and ease of use. Most users recommended its integration into coursework.

Despite being early-stage, users welcomed the tool's potential. Future improvements include support for biaxial simulations and custom structural models. Planned research will evaluate long-term learning outcomes and compare results with a control group using traditional materials. Broader testing across diverse academic settings is also planned to validate the tool's effectiveness and adaptability.

6 Acknowledgement

We gratefully acknowledge the use of the RWTHDynLab at RWTH Aachen University, where the shaking table experiments were conducted. The shaking table was funded by the German Research Foundation (DFG) under grant numbers INST 222/1161-1 FUGG and INST 222/1403-1 FUGG.

7 Conflict of Interest

The authors declare that this research was carried out without any commercial or financial interests that could be perceived as a potential conflict of interest.

References

[1] P. X. Candeias, A. A. Correia, and G. G. Tekeste, "Shake table testing techniques: Current challenges and new trends", in *Advances on Testing and Experimentation in Civil Engineering: Materials, Structures and Buildings*, C. Chastre, J. Neves, D. Ribeiro, M. G. Neves, and P. Faria, editors. Cham: Springer International Publishing, 2023, pp. 173–197. DOI: 10.1007/978-3-031-23888-8_8

[2] N. Jovanović, "Exploring the seismic behavior of buildings through shake table testing", *World J. Adv. Eng. Technol. Sci.*, vol. 14, no. 1, pp. 020–030, Jan. 2025.

[3] G. Balaskas, N. Lenzen, F. Kolisch, B. Hoffmeister, and S. Klinkel, *Dataset on the experimental investigation of the seismic response of moment-resisting steel frames using shaking table tests*, 2024.

[4] J. Grashorn, M. Bittner, M. Banse, X. Chang, M. Beer, and A. Fau, "Namazu: Low-cost tunable shaking table for vibration experiments under generic signals", en, *Exp. Tech.*, vol. 49, no. 1, pp. 97–115, Feb. 2025.

[5] M. Morano et al., "Shake table testing of a full-scale three-story seismically resilient steel-framed building with reconfigurable lateral force-resisting systems", en, *Earthq. Spectra*, vol. 41, no. 2, pp. 1389–1420, May 2025.

[6] E. Damcı and Ç. Şekerci, "Development of a low-cost single-axis shake table based on arduino", en, *Exp. Tech.*, vol. 43, no. 2, pp. 179–198, Apr. 2019.

[7] M. S. Corporation, *MTS 3.0 Meter Biaxial Seismic Simulator*, https://www.mts.com/en/products/civil-engineering/seismic-simulators.

[8] M. Conrad, D. Kablitz, and S. Schumann, "Learning effectiveness of immersive virtual reality in education and training: A systematic review of findings", en, *Computers & Education: X Reality*, vol. 4, no. 100053, p. 100 053, 2024.

[9] K. H. James, G. K. Humphrey, T. Vilis, B. Corrie, R. Baddour, and M. A. Goodale, ""active" and "passive" learning of three-dimensional object structure within an immersive virtual reality environment", *Behav. Res. Methods Instrum. Comput.*, vol. 34, no. 3, pp. 383–390, Aug. 2002.

[10] Z. Chen et al., "Shakebot: A low-cost, open-source robotic shake table for earthquake research and education", 2022. eprint: 2212.10763 (cs.RO).

[11] T. Yamashita, M. K. Pal, K. Matsuzaki, and H. Tomozawa, "Development of a virtual reality experience system for interior damage due to an earthquake – utilizing E-Defense shake table test –", en, *J. Disaster Res.*, vol. 12, no. 5, pp. 882–890, Oct. 2017.

[12] T. Al-Zuriqat, P. P. Abadia, C. C. Geck, K. Dragos, and K. Smarsly, "Implementation and validation of a low-cost lot-enabled shake table system", in *Proceedings of the 14th International Workshop on Structural Health Monitoring*, Destech Publications, Inc., Sep. 2023.

[13] D. Isobe and Q. Yang, "An integrated finite element analysis and virtual reality system for structural and indoor nonstructural components of buildings under seismic excitations", en, *J. Build. Eng.*, vol. 98, no. 111320, p. 111 320, Dec. 2024.

[14] R. L. Chandra, D. Berkaoui, P. Querl, K. Castermans, and H. Nacken, "Student perceptions of VR labs modeled after traditional physical labs in non-destructive measurement testing", *Int. J. Inf. Educ. Technol.*, vol. 15, no. 3, pp. 419–427, 2025.

Parametrische Simulation von Linienbaustellen für KI-gestützte Ressourcenplanung

Anna-Maria Mehringer[1,2], Jonas Wiederer[1], Maik Görgner[1,2] and Simon Höng[1]
[1]Ostbayerische Technische Hochschule Regensburg, Seybothstraße 2, 93053 Regensburg, Germany
[2]Firmengruppe Max Bögl, Max-Bögl-Straße 1, 92369 Sengenthal, Germany
E-mail(s): anna-maria.mehringer@oth-regensburg.de, jonas.wiederer@oth-regensburg.de, maik.goergner@oth-regensburg.de, simon.hoeng@oth-regensburg.de

Abstract: Diese Arbeit präsentiert ein Konzept zur automatisierten Generierung und Bewertung von Entwurfsvarianten für Linienbaustellen. Basierend auf definierten Rahmenbedingungen und Zielvorgaben werden mittels eines Algorithmus konsistente Simulationsparameter abgeleitet und in ein parametrisches Simulationsmodell überführt. Die Simulation bildet den Bauablauf einer Erdbau-Linienbaustelle ab und ermittelt zentrale Kennwerte wie Bauzeit, Standzeiten, transportierte Materialvolumina und Fahrstrecken. Durch systematische Variation der Einflussgrößen, insbesondere der Fahrzeugflotte, entstehen umfangreiche Datensätze, die sowohl für das Training einer Künstlichen Intelligenz (KI) als auch für die iterative Optimierung verschiedener Bauvarianten eingesetzt werden.

Die Ergebnisse zeigen, dass die entwickelte Simulation eine wirksame Optimierung der Transportlogistik ermöglicht. Dabei wird deutlich, dass verschiedene Parameterkonfigurationen zu ähnlichen Ergebnissen führen können und die Gewichtung der Optimierungsziele entscheidend ist. Die Validierung mittels evolutionärem Algorithmus bestätigt die Funktionsfähigkeit des Systems und zeigt, dass eine ausschließlich bauzeitbasierte Bewertung eindimensional ist.

Die simulierten Varianten sollen hinsichtlich technischer, wirtschaftlicher und ökologischer Kriterien bewertet werden. Die konkrete Umsetzung erfolgt in *Tecnomatix Plant Simulation* und bildet die Grundlage für ein automatisiertes Generative-Design-System zur Optimierung von Linienbaustellen.

Keywords: Generative Design, Construction Simulation, AI Optimization, Linear Construction

DOI: 10.18154/RWTH-CONV-254890. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Motivation

Bis 2030 sollen 80 Prozent des Bruttostromverbrauchs in Deutschland durch erneuerbare Energien abgedeckt werden [1]. Dies erfordert einen umfassenden Ausbau der Energieinfrastruktur, insbesondere des Übertragungsnetzes, um die im Norden erzeugte Windenergie in die südlichen Verbrauchszentren zu transportieren. Hierfür kommen Hochspannungsstromtrassen zum Einsatz. Das Bundesbedarfsplangesetz fördert seit 2013 den Bau von 35 Vorhaben mit rund 14.000 Kilometer Leitungslänge.

Bis 2023 waren jedoch nur etwa 4.000 Kilometer genehmigt oder fertiggestellt [2]. Diese Diskrepanz verdeutlicht den Bedarf nach effizienter Planung und einem effektiven Projektmanagement.

Aufgrund der Neuartigkeit und Dimension dieser Bauvorhaben fehlen belastbare Erfahrungswerte und standardisierte Verfahren zur Effizienzsteigerung. Vorhandenes Wissen bleibt oft an einzelne Personen gebunden und geht bei Projektende oder Personalwechsel verloren. Die mangelnde Dokumentation erschwert die systematische Übertragung bewährter Praktiken auf neue Vorhaben. Ein ganzheitlicher, KI-gestützter Planungsansatz bietet hier Potenzial: Ein automatisierter Workflow kann aus definierten Rahmenbedingungen mittels Generative Design verschiedene Entwurfsvarianten generieren und parametrisch simulieren. Diese werden hinsichtlich Bauzeit, Kosten, Ressourceneinsatz, Logistik und Nachhaltigkeit bewertet und optimiert. Dies erleichtert die Entscheidungsfindung erheblich, da Planende aus verschiedenen Varianten mit entsprechenden Leistungskennzahlen wählen können, anstatt diese selbst aufwendig berechnen zu müssen.

2 Stand der Technik

Die Anwendung von Simulationstechniken zur Optimierung von Erdbauprozessen wurde bereits 2010 im Forschungsprojekt ForBAU - Virtuelle Baustelle untersucht. Wimmer et al. erforschten mittels diskreter ereignisorientierter Simulationen (DES) die Verbesserung von Maschineneinsatz und Logistikplanung von Erdbaustellen [3]. Hierbei wurden Trasse, Baustellenlayout, Projektplan und Maschinendaten als Eingangsdaten genutzt. Das Tool dient als Entscheidungshilfe für die optimale Einteilung wesentlicher Ressourcen. Bereits damals wurde eine Notwendigkeit in der Einbeziehung einer Nachhaltigkeitsberechnung gesehen. Mit diesem Ansatz beschäftigte sich Hovanec et al. und entwickelte einen Simulationsansatz für Life-Cycle-Assessment-Analysen (LCA) von Fahrzeugen [4].

DES bieten einen vielversprechenden Ansatz zur Generierung synthetischer Daten, wodurch eine saubere Datengrundlage für das KI-Training geschaffen werden kann [5]. Auch im Bauwesen zeigten Höng et al. die Erzeugung synthetischer Daten aus parametrischen BIM-Modellen und trainierten damit erfolgreich ein KI-System zur Objektklassifizierung [6]. Behl et al. demonstrierten, dass durch optimierte Simulationsparameter qualitativ hochwertige synthetische Datensätze entstehen, die den Trainingsaufwand reduzieren und die Modellgenauigkeit verbessern [7].

Während Wimmer et al. das Baustellenlayout mittels mathematischer Optimierungsverfahren zur Minimierung der Transportwege als definierten Eingangswert der Simulation festlegte, bietet die Kombination von selbstlernenden KI-Algorithmen und Simulationen die Möglichkeit, die Baustelleneinrichtung dynamisch zu optimieren. Diese innovative Herangehensweise wurde von Choi et al. sowie Klar et al. erfolgreich für Fertigungsbetriebe demonstriert [8] [9].

Wiederer et al. konzipierten einen Optimierungsworkflow für die Brückenbauwerksplanung, der vollparametrische Brückenmodelle über automatisierte Schnittstellen mit Bewertungsprozessen verknüpft [10]. Durch kontinuierliche Rückgabe von Kenngrößen entsteht eine Optimierungsschleife, die manuell oder durch KI-Algorithmen gesteuert werden kann.

Die vorliegende Arbeit grenzt sich durch die Entwicklung eines systemübergreifenden Optimierungsworkflows ab, der Bauwerksanalyse mit weiterführenden Analysen zur Quantifizierung multipler Faktoren verknüpft. Im Gegensatz zu bisherigen Arbeiten, die sich primär auf einzelne Optimierungsaspekte

fokussieren, wird hier die Simulation als zentraler Ausgangspunkt für alle relevanten Planungsparameter eingesetzt. Die Besonderheit liegt in der Integration einer KI, die auf diesem systemübergreifenden Workflow trainiert werden soll und dadurch eine kontinuierliche Optimierungsschleife ermöglicht. Diese neuartige Methodik fokussiert sich auf die Optimierung von Energietrassenbaustellen, wobei die theoretische Grundlage für die Anwendbarkeit von Simulationen zur Ressourcenoptimierung die Untersuchung von Daubenmerkl zu Erdaushubprozessen liefert [11].

3 Methodik

Die Realisierung großmaßstäblicher Trassenprojekte stellt hohe Anforderungen an Planungsgenauigkeit, Ressourceneffizienz und Terminsicherheit. Die vorliegende Arbeit präsentiert ein softwaregestütztes Konzept, das mithilfe Künstlicher Intelligenz automatisch Entwurfsvarianten für den Ressourceneinsatz und den Bauablauf entlang linearer Trassen erstellen soll. Ziel ist die automatisierte Überführung unterschiedlicher Randbedingungen in ressourcentechnisch bewertbare Ausführungsvarianten unter Berücksichtigung von Leistungskennzahlen.

Die Quantifizierung umfasst Transportrouten, Lagerplätze sowie Leistungskennzahlen zu Kosten, Flottenmanagement und grauer Energie. Als Grundlage fungiert eine parametrische Simulation in *Tecnomatix Plant Simulation (TPS)*, deren strukturierter Kennwertexport die nachgelagerte Bewertung von Bauablaufsvarianten ermöglicht. Integrierte KI-Optimierungsalgorithmen ergänzen den Workflow und ermöglichen eine automatisierte Parameteroptimierung zur Erfüllung vordefinierter Zielvorgaben.

3.1 Konzept

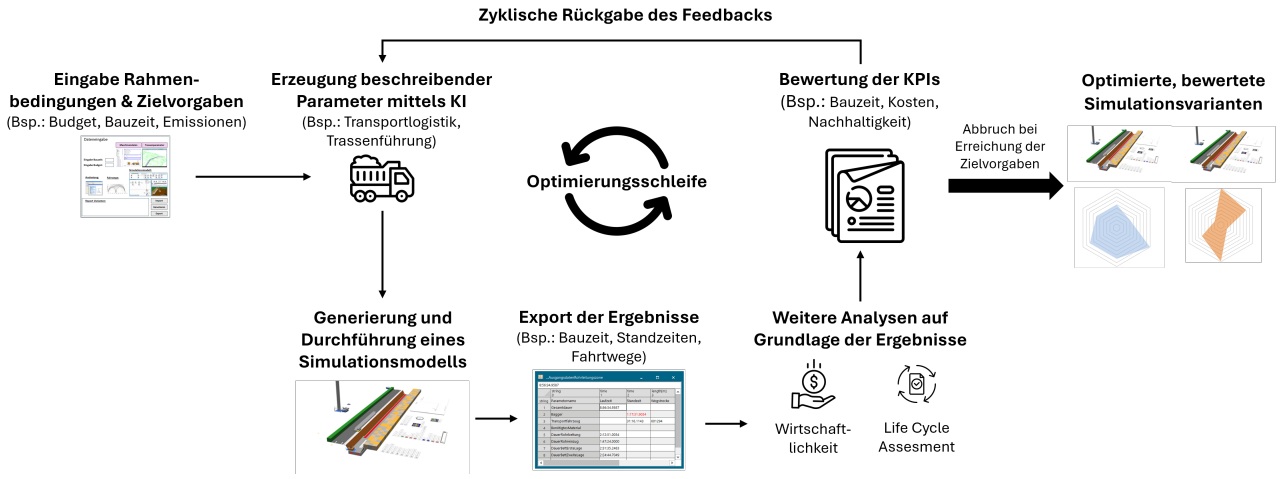


Abbildung 1: Zirkulärer Optimierungsworkflow für simulationsbasierte Linienbaustellen

Ausgehend von den Rahmenbedingungen einer Linienbaustelle werden verschiedene Bauvarianten automatisch generiert und hinsichtlich ihrer Bauzeit, Kosten, Ressourceneinsatz, Logistik und Nachhaltigkeit bewertet sowie optimiert. Als Grundlage für den Workflow dient eine parametrische Simulation, die den Bauablauf einer Linienbaustelle abbildet. Zu Beginn werden alle relevanten Rahmenbedingungen und Zielgrößen durch den Nutzer definiert. Anschließend wertet ein Algorithmus diese Vorgaben aus und leitet konkrete Simulationsparameter ab, wie die Auswahl bestimmter Baumaschinen, deren Anzahl oder die Entfernung zu Zwischenlagern. Auf Basis dieser Daten entsteht ein projektspezifisches

Simulationsmodell, das zentrale Kennzahlen wie die Gesamtbauzeit, die Standzeiten der Maschinen und die gefahrenen Strecken ermittelt. Die ermittelten Kennzahlen können in weiterführende Planungstools integriert werden, sodass der zeitliche Aspekt um zusätzliche ökologische und wirtschaftliche Gesichtspunkte ergänzt wird. Daraufhin erfolgt eine Bewertung basierend auf den Zielvorgaben des Users. Die Ergebnisse fließen in eine Optimierungsschleife zurück, in der der Algorithmus die Parameter schrittweise anpasst, bis die definierten Zielfunktionen erreicht sind. Auf diese Weise entsteht ein iteratives Verfahren, das technische, wirtschaftliche und ökologische Zielgrößen in Einklang bringt.

3.2 Parametrisches Simulationsmodell für ressourcenbezogene Trassenprozesse

Das parametrische Simulationsmodell in *TPS* bildet die typischen Prozesse der Erdbaulogistik entlang von Energietrassenbaustellen ab. Dazu zählen Aushub, Transport, Lagerung und Einbau des Erdmaterials und der Leitungen unter Berücksichtigung von Maschineneinsatz, Geometrie und Schichtenaufbau. Als fixierte Steuergrößen dienen vom Bauherren vorgegebene Projektparameter wie das Grabenprofil, die Bodenbeschaffenheit, die Abschnittslängen und Lagerflächen. Die Eingabe soll über eine intuitive grafische Benutzeroberfläche erfolgen. Durch die gezielte Variation von Ressourcen- und Logistikparametern (z. B. Maschinenanzahl, Transportdistanzen, Ladekapazitäten) soll der zeitlich optimale Bauablauf bestimmt werden.

3.3 Automatisierte Pipeline zur Berechnung von weiteren Kenngrößen

Auf Grundlage der Simulationsergebnisse sollen mittels einer automatisierten Pipeline kostenrelevante Kennwerte, Ressourcenzuordnungen, ökologische Indikatoren sowie Terminplanungen systematisch ermittelt werden. Für die Übergabe der Simulationsergebnisse können entweder native *TPS*-Schnittstellen, beispielsweise Datenbanken, oder der Export im JSON-Format genutzt werden. Dadurch liegt für jede Variante eine fundierte Bewertung vor, die zugleich als Trainingsgrundlage für die Entwicklung erdbauspezifischer Ressourcensimulationsmodelle dienen kann.

3.4 Erzeugung der Simulationsparameter mittels KI

Die entwickelte Simulation sowie die darauf aufbauenden Analysen arbeiten deterministisch und vorwärtsgerichtet. Das bedeutet, dass auf Basis definierter Eingabeparameter (z. B. Trassen- und Bodenkennwerte, Logistikparameter) die entsprechenden Ergebnisgrößen (z. B. Bauzeit, Kosten, Nachhaltigkeit) berechnet werden. Im praktischen Baukontext stehen jedoch nicht primär die Eingabeparameter, sondern die angestrebten Zielgrößen im Vordergrund. Hieraus ergibt sich eine inverse Problemstellung, welche Entwurfsparameter zu einer vorgegebenen Zielsetzung führen (beispielsweise die minimale Bauzeit bei einem festgelegten Budget). Zur Lösung dieser Problemstellung werden die Simulationsmodelle in Kombination mit den weiterführenden Analysen zunächst zur Generierung umfangreicher synthetischer Datensätze eingesetzt, welche Eingabeparameter mit den zugehörigen Ergebnisgrößen verknüpfen. Auf dieser Grundlage soll mittels überwachten Lernens ein KI-System trainiert werden, das auf Basis projektspezifischer Daten (z. B. Trassenverlauf, baustellenspezifische Rahmenbedingungen) und definierter Zielvorgaben (z. B. Budget, Bauzeit) optimale Ressourcen- und Logistikparameter ableitet. Ziel ist es, unter Einhaltung der vorgegebenen Randbedingungen einen möglichst effizienten und störungsarmen Bauablauf zu gewährleisten.

3.5 Rückgabe der Bewertung

Die ermittelten Ergebnisgrößen werden nun anhand der vom Nutzer vorab definierten Zielvorgaben evaluiert. Bei unvollständiger Zielerfüllung erfolgt ein Feedback an das KI-System, welches daraufhin neue Entwurfsparameter generiert. Auf Basis dieser Parameter wird der Simulations- und Bewertungsprozess erneut gestartet. Treten Zielkonflikte zwischen verschiedenen Optimierungszielen auf, werden diese dem Nutzer in geeigneter visueller Form aufbereitet, um eine fundierte Entscheidungsfindung zu ermöglichen. Auf diese Weise entsteht eine iterative Optimierungsschleife, die so lange durchlaufen wird, bis entweder eine definierte Anzahl an Varianten vorliegt, welche sämtliche Zielvorgaben erfüllen, oder eine festgelegte Anzahl an Varianten ohne signifikante Verbesserung erzeugt wurde.

4 Anwendungsbeispiel: Simulation einer konkreten Erdbaustelle

Die Simulation bildet das fundamentale Element des vorgestellten KI-gestützten Workflows und muss deshalb die Bauprozesse möglichst korrekt darstellen. Das entwickelte Modell rekonstruiert die offene Bauweise von Energietrassen, bei der Hochspannungskabel in offenen Gräben verlegt werden. Die nachgelagerten KI-Optimierungsalgorithmen und die integrierte Bewertungslogik stellen derzeit noch konzeptionelle Ansätze dar, deren Implementierung auf der entwickelten Simulation aufbaut.

4.1 Aufbau der Simulationen

Das konzeptionelle Framework der Simulation ist in Abbildung 2 visualisiert. Zur Modellentwicklung wurde ein Bottom-Up-Ansatz gewählt, um detaillierte Kennwerte einzelner Prozessschritte zu ermitteln und diese modular in ein Gesamtmodell zu integrieren. Als technische Basis fungiert ein definierter Regelquerschnitt des Grabens, welcher die Grundlage für sämtliche nachgelagerte Simulationsmodelle bildet. Der Bauablauf der offenen Bauweise wird in fünf sequenzielle Teilprozesse untergliedert: Oberbodenabtrag, Grabenaushub, Herstellung der Rohrleitungszone, Rückverfüllung und Oberbodenandeckung.

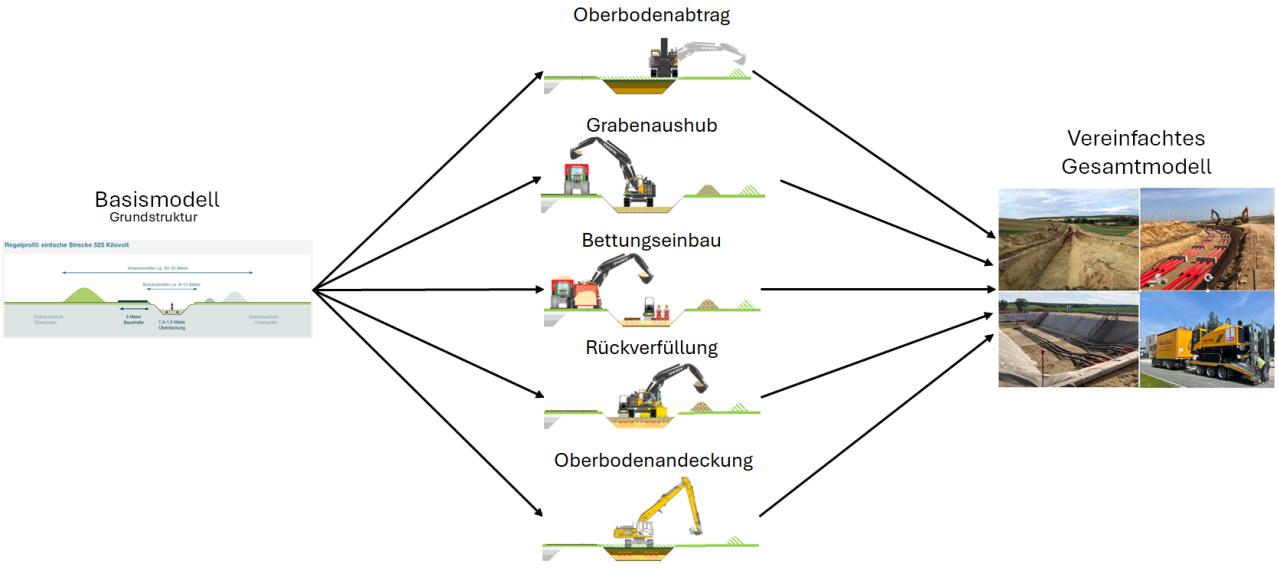


Abbildung 2: Konzeptioneller Aufbau des modularen Simulationssystems nach dem Bottom-Up-Ansatz

Jeder Teilprozess wird mittels separatem Simulationsmodell analysiert. Die zeitliche Grundlage der Ermittlung bilden empirische Aufwandswerte aus realen Kalkulationsgrundlagen, während die Parametrisierung durch projektspezifische Parameter zu Baustellenbedingungen und Ressourceneinsatz erfolgt. Für definierte Abschnittslängen werden Bau- und Standzeiten, bewegte Erdvolumina sowie Transportstrecken ermittelt. Die Integration der Einzelprozesse erfolgt im übergeordneten Modell durch sequenzielle Verknüpfung. Der analytische Fokus liegt auf der zeitlichen Taktung der Arbeitsschritte sowie der Identifikation potenzieller Prozesskollisionen. Das Gesamtsystem ermöglicht die Bestimmung des optimalen Maschineneinsatzes hinsichtlich Anzahl, zeitlichem Einsatz und räumlicher Verteilung zur Gewährleistung eines kontinuierlichen Baufortschritts. Als Eingangsdaten dienen die aus den Teilprozessen ermittelten Kenngrößen. Diese Methodik ermöglicht die Ableitung realistischer, parameterabhängiger Aufwandswerte zur präzisen Bestimmung der Gesamtbauzeit.

Zur Vermeidung von Schnittstellenproblemen sind alle Simulationsbausteine in einer gemeinsamen Datei organisiert und fungieren konzeptionell als eigenständige Module. Diese modulare Struktur ermöglicht die gezielte Analyse einzelner Arbeitsschritte sowie eine flexible Anpassung des Modells bei veränderten Randbedingungen.

4.2 Durchführung von Versuchsreihen

Zur Anwendung und Validierung des entwickelten Simulationsmodells wurden verschiedene Versuchsreihen durchgeführt, um den Einfluss der Parameter auf die Simulationsergebnisse mit realen Daten zu vergleichen. Die Versuchsreihen wurden mit dem softwareinternen Experimentmanager durchgeführt, welcher die Analyse mehrerer zufallsbasierter Simulationsläufe bei festgelegten Parametereinstellungen ermöglicht. Die Ergebnisse werden durch Mittelwertbildung aus einer bestimmten Anzahl von Simulationsläufen ausgewertet.

Für die Prozesse, die unabhängig von Materialtransporten sind (Oberbodenabtrag und -andeckung, Rückverfüllung), entsteht durch den zufallsverteilten Einfluss von Störungen eine Zeitspanne, die sich mit der Angabe der Tagesleistung deckt. Bei den übrigen Prozessen wurde die Anzahl, die Ladekapazität und die Taktung der Transportfahrzeuge untersucht. Die Erhöhung von einem auf mehrere Fahrzeuge bewirkt eine signifikante Leistungssteigerung. Bei weiterer Erhöhung der Fahrzeuganzahl zeigt sich lediglich eine marginale Effizienzsteigerung bei gleichzeitig ansteigenden Standzeiten der Transportfahrzeuge. Ab einer Fahrzeuganzahl von zwei Einheiten führt eine Erhöhung der Ladekapazität zu vernachlässigbaren Leistungsverbesserungen. Diese Beobachtung bestätigt die Erkenntnisse von Wimmer [3]. Bezüglich der Fahrzeugtaktung erwies sich der minimal eingesetzte Wert als optimal, da sich die Fahrzeuge bei der gegebenen Transportentfernung ab einem kritischen Punkt einholen, wodurch die Taktung ihre steuernde Wirkung verliert.

Zur weiteren Validierung der Simulationsergebnisse wurde eine Teilsimulation der Rohrleitungszone mittels evolutionärem Algorithmus optimiert. Der betrachtete Trassenabschnitt betrug 48 Meter bei einer mittleren Entfernung von 25 Kilometer zum Materiallager. Die Transportgeschwindigkeiten werden mit Schrittgeschwindigkeit auf der Baustraße und durchschnittlich 60 km/h auf öffentlichen Straßen angesetzt.

Die multikriteriellen Optimierungsparameter sind in Tabelle 1 dargestellt. Optimiert wurden in absteigender Priorität: Gesamtbauzeit, Standzeit des Baggers und Standzeiten der Transportfahrzeuge.

Tabelle 1: Parameterbereiche und Schrittweiten für die genetische Optimierung

Optimierungsparameter	Untere Stufe	Obere Stufe	Schrittweite
Anzahl der Fahrzeuge	1	3	1
Ladekapazität	7	12	nicht linear (7,9,12)
Taktung	30 min	2 h	30 min

Der evolutionäre Algorithmus bestimmte als optimale Konfiguration drei Transportfahrzeuge mit 9 m³ Ladekapazität bei 30-minütiger Taktung. Diese Lösung steht scheinbar im Widerspruch zu den vorangegangenen Erkenntnissen über marginale Effizienzsteigerungen bei mehr als zwei Fahrzeugen und höherer Ladekapazität. Das Ergebnis resultiert jedoch aus der 4:1-Gewichtung der Bagger-Standzeit gegenüber den Transporter-Standzeiten. Ein Kontrolllauf mit gleichgewichteten Standzeiten bestätigte den starken Einfluss der Gewichtung. Die optimale Konfiguration reduzierte sich auf zwei Fahrzeuge mit 12 m³ Ladekapazität bei unveränderter Gesamtbauzeit, jedoch bei halbierter Standzeit der Transportfahrzeuge und bestätigte damit die zuvor ermittelten Zusammenhänge.

5 Diskussion und Fazit

Die Ergebnisse verdeutlichen, dass die entwickelte Simulation eine effektive Optimierung der Transportlogistik ermöglicht. Es zeigt sich, dass unterschiedliche Parameterkonfigurationen zu vergleichbaren Ergebnisgrößen führen können, wobei die Gewichtung der einzelnen Optimierungsziele einen maßgeblichen Einfluss auf das Resultat hat. Das Beispiel macht zudem deutlich, dass eine ausschließlich auf die Bauzeit ausgerichtete Bewertung zu einer eindimensionalen Betrachtung führen kann, da bei identischer Bauzeit dennoch signifikante Unterschiede in den zugrunde liegenden Prozessen bestehen können. Daher ist eine Erweiterung der Bewertungsgrundlage um Kosten- sowie ökologische Kenngrößen erforderlich, um eine ganzheitliche und belastbare Entscheidungsgrundlage zu schaffen.

Die entwickelte parametrische Simulation bildet die Grundlage für ein automatisiertes Generative-Design-System zur Optimierung von Energietrassen. Das modulare Simulationsmodell ist in der Lage, den gesamten Bauablauf der offenen Bauweise abzubilden und liefert zentrale Kennzahlen wie Bauzeit und Ressourceneinsatz, welche als Eingabegrößen für nachgelagerte Bewertungs- und Optimierungsalgorithmen dienen. Durch den parametrischen Grabenquerschnitt können zwar unterschiedliche Grabenprofile berücksichtigt werden, jedoch ist für eine realitätsnahe Abbildung die Integration weiterer Faktoren erforderlich, wie beispielsweise der standortspezifischen Bodenverhältnisse sowie trassenspezifischer Kennwerte. Auch im Gesamtmodell der Simulation muss die Logik der Baggerabfolge und der Parallelisierung von Tätigkeiten noch so modelliert werden, dass ein konsistenter und praxisgerechter Bauablauf entsteht.

Der gegenwärtige Entwicklungsstand umfasst die Implementierung der Basissimulation in *TPS*. Die nächsten Entwicklungsschritte konzentrieren sich auf die Weiterentwicklung der Simulation sowie die Implementierung der automatisierten Pipeline zur systematischen Überführung der Simulationsergebnisse in kostentechnische, ressourcenbezogene und nachhaltigkeitsorientierte Metriken. Darüber hinaus sind die automatisierte Parameterableitung mittels KI-Algorithmen sowie die Implementierung einer iterativen Optimierungsschleife vorgesehen.

Acknowledgements

Diese Veröffentlichung wurde im Zuge der Masterarbeit durch die Firmengruppe Max Bögl gefördert.

Literatur

[1] *Erneuerbare-Energien-Gesetz (EEG 2023)*, 2023. besucht am 2. Juli 2025. Adresse: https://www.gesetze-im-internet.de/eeg_2014/

[2] »Netzausbau 2024: Aktuelle Zahlen und Entwicklungen«, besucht am 2. Juli 2025. Adresse: https://www.enercity.de/magazin/unsere-welt/warum-das-stromnetz-ausgebaut-werden-muss

[3] J. Wimmer, T. Horenburg, W. A. Günthner, Y. Ji, A. Borrmann und E. Rank, »Evaluierung einer 3D-Modell-basierten Ablaufsimulation von Erdbauprozessen in der Praxis«, in *Integrationsaspekte der Simulation: Technik, Organisation und Personal*, G. Zülch und P. Stock, Hrsg., KIT Scientific Publishing, 2010, S. 142–148.

[4] M. Hovanec, P. Korba, V. Tymofiiv, I. Sekelová und L. Főző, »Simulating Decarbonization of the Road Transport Sector by Siemens Plant Simulation«, in *SISY 2023 – 21st International Symposium on Intelligent Systems and Informatics*, Pula, Croatia: IEEE, 2023, S. 377–382. DOI: 10.1109/SISY60376.2023.10417907

[5] K. C. Chan, M. Rabaev und H. Pratama, »Generation of synthetic manufacturing datasets for machine learning using discrete-event simulation«, *Production & Manufacturing Research*, Jg. 10, Nr. 1, S. 337–353, 2022. DOI: 10.1080/21693277.2022.2086642

[6] S. K. Höng und F. A. Eder, »Methode zur Generierung multimodaler synthetischer Daten aus parametrischen BIM-Modellen zur Nutzung in KI-Systemen«, in *Tagungsband 34. Forum Bauinformatik 2023, Bochum, Deutschland*, 2023. DOI: 10.13154/294-10130

[7] H. S. Behl, A. G. Baydin, R. Gal, P. H. Torr und V. Vineet, »AutoSimulate: (Quickly) Learning Synthetic Data Generation«, in *Lecture Notes in Computer Science*, Springer, 2020. DOI: 10.1007/978-3-030-58542-6_16 Adresse: https://www.researchgate.net/publication/346987788

[8] H. Choi u. a., »Optimization of the Factory Layout and Production Flow Using Production-Simulation-Based Reinforcement Learning«, *Machines*, Jg. 12, S. 390, 2024. DOI: 10.3390/machines12060390

[9] M. Klar u. a., »Transferable Multi-Objective Factory Layout Planning Using Simulation-Based Deep Reinforcement Learning«, *Journal of Manufacturing Systems*, Jg. 74, S. 487–511, 2024. DOI: 10.1016/j.jmsy.2024.04.007

[10] J. Wiederer, S. K. Höng und F. A. Eder, »Konzept zur KI-gestützten parametrischen Brückenmodellierung für ressourcen- und kostenoptimierte Bauwerksentwürfe«, in *Tagungsband 35. Forum Bauinformatik 2024, Hamburg, Deutschland*, 2024. DOI: 10.15480/882.13506

[11] K. Daubenmerkl, »Simulation einer Linienbaustelle - Virtuelle Ressourcenoptimierung von Erdbauprozessen am Beispiel der Energietrasse „SuedLink“«, Magisterarb., OTH Regensburg, 2024.

Graphbasierte Visualisierung von Produktdaten der Technischen Gebäudeausrüstung

Erich Dierkes[1]
[1]Lehr- und Forschungsgebiet für Computergestütztes Entwerfen (DC), RWTH Aachen University, Aachen, Deutschland
E-mail(s): dierkes@dc.rwth-aachen.de

Abstract: Die Digitalisierung der Technischen Gebäudeausrüstung (TGA) erfordert standardisierte, maschinenlesbare Produktdaten. Obwohl DIN EN ISO 16757 einen internationalen Standard vorgibt, liegen Herstellerkataloge häufig weiterhin in VDI-3805-Strukturen vor. Der Beitrag adressiert diese Lücke durch eine systematische Analyse von VDI 3805-1 (Grundlagen) und VDI 3805-22 (Wärmepumpen) sowie durch eine prototypische, graphbasierte Aufbereitung der Daten. Methodisch werden Satzarten und Hierarchien formal modelliert und validiert; eine interaktive Graphvisualisierung macht Klassifikation, Berechnungsfunktionen, Zubehörabhängigkeiten und Geometrie transparent. Als Anwendungsfall dient ein Wärmepumpen-Datensatz. Die Untersuchung identifiziert wiederkehrende Struktur- und Abhängigkeitsmuster und erläutert die Rolle der TGA-Nummer bei Variantenführung und Referenzen. Sie zeigt außerdem, wie Funktionen virtuelle Datensätze erzeugen und dynamische Merkmale parametrisieren. Der Prototyp hebt Regelverweise hervor, deckt Inkonsistenzen und unvollständige Relationen frühzeitig auf und beschleunigt die Navigation sowie das Mapping relevanter Merkmale. Damit entsteht eine belastbare Grundlage für die Transformation nach DIN EN ISO 16757, inklusive Vorstrukturierung von Datenwörterbüchern und Ableitung von Austauschobjekten.

Keywords: Produktdaten, TGA, VDI3805, DIN EN ISO 16757, IFC

DOI: 10.18154/RWTH-CONV-254880. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Einleitung

Die Digitalisierung der technischen Gebäudeausrüstung erfordert interoperable, strukturierte Produktdaten. Trotz der internationalen Norm DIN EN ISO 16757 dominieren in vielen Herstellerkatalogen weiterhin VDI-3805-Strukturen. VDI 3805 Blatt 1 stellt hierfür eine einheitliche Datenorganisation bereit (technische, geometrische, mediale Aspekte) [1]. Ein tiefes Verständnis dieser Datenstrukturen ist Voraussetzung für Migration, Automatisierung und eine belastbare Einbindung in BIM-Workflows.

Dieser Beitrag berichtet aus einem laufenden Forschungsprojekt zur Übersetzung von VDI 3805 in DIN EN ISO 16757. Im Fokus stehen die Analyse und Aufbereitung der Richtlinien VDI 3805-1 (Grundlagen) [1] und VDI 3805-22 (Wärmepumpen) [2]. Die zentrale Forschungsfrage lautet: Wie können VDI 3805-Strukturen systematisch analysiert und visualisiert werden, um eine fundierte Grundlage für die Transformation zu DIN EN ISO 16757 zu schaffen? Ziel ist es, Datenstruktur und

Modellierungslogik der VDI-Kataloge zu verstehen, zu dokumentieren und mittels graphbasierter Visualisierung zugänglich zu machen; die daraus gewonnenen Erkenntnisse bilden die Basis für spätere Modelltransformation und Konfiguration im Kontext von DIN EN ISO 16757.

Methodisch werden die Produktdaten formal mit *Pydantic* beschrieben und validiert, um Satzarten und Hierarchien maschinenlesbar und prüfbar abzubilden. Für die explorative Darstellung der Beziehungen wird ein interaktiver Graph-Viewer auf Basis von *Dash Cytoscape* eingesetzt, der die semantische Struktur der Daten veranschaulicht. Als exemplarische Anwendung dient ein Datensatz nach VDI 3805, der die hierarchische Struktur und die logischen Beziehungen innerhalb eines Wärmepumpenprodukts beschreibt.

2 Stand der Technik

Die Planung, Konstruktion, Berechnung und Simulation im Bereich der Technischen Gebäudeausrüstung (TGA) erfordert zunehmend den Austausch von Produktdaten in maschinenlesbarer Form. Um die Effizienz in der rechnergestützten Planung zu steigern und eine einheitliche Datenverarbeitung zu gewährleisten, wurden verschiedene Normen und Richtlinien etabliert. Diese zielen darauf ab, herstellerspezifische Produktinformationen, die traditionell in Katalogen mit Tabellen, Diagrammen und Zeichnungen vorlagen, in strukturierter und interoperabler Form bereitzustellen. Zwei zentrale Normenwerke prägen den aktuellen Stand der Technik im Bereich des Produktdatenaustauschs für TGA: die DIN EN ISO 16757-Reihe und die VDI 3805-Reihe.

2.1 DIN EN ISO 16757 - Datenstrukturen für elektronische Produktkataloge der Technischen Gebäudeausrüstung

Die Normenreihe DIN EN ISO 16757 legt Datenstrukturen für elektronische Produktkataloge der Technischen Gebäudeausrüstung fest [3]. Sie wurde vom Technischen Komitee ISO/TC 59 „Buildings and civil engineering works“ der Internationalen Organisation für Normung (ISO) erarbeitet und durch das Technische Komitee CEN/TC 442 „Building information modelling (BIM)“ als Europäische Norm übernommen.

Die Normenreihe (aktuell u.a. Teil 1 Konzepte [3], Teil 2 Geometrie [4]) definiert ein generisches Referenzmodell für Produktkataloge und deren Geometriedaten. Entwürfe für Skriptsprache, Datenwörterbücher und Austauschformat (Teile 3-5) liegen vor [5], [6]. Kernziel ist die interoperable, automatische Nutzung herstellerneutral beschriebener Produkt- und Geometriedaten in BIM-Werkzeugen.

2.2 VDI 3805 - Produktdatenaustausch in der Technischen Gebäudeausrüstung

Die VDI 3805-Richtlinienreihe hat sich in Deutschland als etablierter Standard für den digitalen Produktdatenaustausch in der Technischen Gebäudeausrüstung (TGA) etabliert. Ihr übergeordnetes Ziel ist es, die Effizienz in der Planung, Konstruktion, Berechnung und Simulation im TGA-Bereich durch die Bereitstellung maschinenlesbarer Produktdaten deutlich zu steigern. Hierzu standardisiert die VDI 3805 den Austausch herstellerspezifischer Informationen, die früher in gedruckten Katalogen bereitgestellt wurden, und überführt sie in eine strukturierte, digitale Form. Dadurch erhalten gleichartige Produkte eine einheitliche Datenstruktur, was die Weiterverarbeitung in Softwarelösungen erheblich erleichtert.

Trotz dieser etablierten Standards bestehen erhebliche Herausforderungen: VDI 3805 weist als 'statisches, nicht maschinenlesbares Datenformat' den wesentlichen Nachteil auf, dass es "von Modellierungs- und Simulationssoftware nicht individuell und flexibel verarbeitet werden kann"[7]. Gleichzeitig ist VDI 3805 als nationaler Standard begrenzt, weshalb DIN EN ISO 16757 entwickelt wurde, üm die Produktdatenabbildung auf ein internationales Niveau zu heben und die Zugänglichkeit zu erweitern"[7]. Diese Transformation erfordert jedoch ein tiefes Verständnis der komplexen VDI-Datenstrukturen, deren hierarchische Abhängigkeiten und semantische Beziehungen bisher nur schwer zugänglich sind. Hier setzt der vorliegende Beitrag an.

3 Methodik

Das Vorgehen gliedert sich in (1) systematische VDI-Analyse, (2) Pydantic-Modellierung und (3) Dash Cytoscape-Visualisierung. Die Richtlinien VDI 3805-1 und VDI 3805-22 wurden nach Satzart-Mustern und Hierarchiebeziehungen analysiert. Pydantic wurde für Typ-Validierung gewählt, Dash Cytoscape für interaktive Graphdarstellung.

Die VDI 3805 nutzt zeilenorientierte Austauschdateien (Endung `.VDI`) mit Semikolon als Trennzeichen.

3.1 Datenstruktur und Merkmale (VDI 3805 Blatt 1)

Die Produktstruktur basiert auf einer Hierarchie (Hauptgruppen → Merkmalsstufen → Produktelemente). Zentral ist die TGA-Nummer als ein rein EDV-mäßiges Konstrukt; Platzhalter ('???', '000') ermöglichen flexible Tiefe [1]. Funktionen und Algorithmen zur Berechnung von Produkteigenschaften können in den Datensatzarten 600 (Deklarationen) und 610 (Definitionen) angegeben werden. Diese Funktionen ermöglichen die Interpretation und direkte Nutzung der Auslegungsalgorithmen durch Berechnungsprogramme und können auf interne Datensätze zur Datenablage zugreifen [1].

Die Funktionen selbst ermöglichen die Zuweisung von Berechnungsformeln und Algorithmen zur Auslegung und Berechnung des Produkts. Sie unterstützen einfache Kontrollstrukturen (IF, DO WHILE) und Unterprogrammstrukturen (FUNCTION). Weiterhin sind spezifische implizite Dateistrukturfunktionen wie VDIINDEX, VDIAKTUELL, VDIFOLGE, VDIVORHER, VDIERSTER, VDILETZTER zum Zugriff auf die Datensätze vorgesehen [1].

3.2 Geometrie und Zubehör (VDI 3805 Blatt 1)

Die Geometriedaten (Satzarten 970 ff.) beschreiben eine TGA-Einheit für Darstellung, Kollisions- und Anschlussprüfung. Sie umfassen allgemeine Geometriedaten (970); Anschlussdaten (970.01), die Lage, Richtung, Art und Abmessungen eines Anschlusses im Raum sowie die Zugehörigkeit definieren und damit automatisierte Installation und geometrische Prüfung ermöglichen; Gestaltdaten (970.02, 970.03) zur Beschreibung der sichtbaren Produktgestalt mit 3D-Bausteinen wie Quadern, Zylindern und speziellen Kanal- und Rohrformen; Materialdaten (970.04) zur Definition von Oberflächenfarben und -strukturen; Symboldaten (970.05, 970.06) zur Beschreibung der Symbolgeometrie; sowie Raumgestaltdaten (970.11-970.14) für Produkt-, Bedien-, Einbring- und Montageraum.

Die Geometriedaten nutzen lokale Koordinatensysteme und können sich auf andere Bausteine beziehen, um komplexe Formen darzustellen. Zubehör kann über Querverweise (Satzarten XX0.01, 760.01, 920, 930) zugeordnet werden. Die Abhängigkeit von Zubehören wird in einem Abhängigkeitsbaum

beschrieben (Satzarten 930 bis 930.04), der auch als Auswahlbaum dienen kann. Bedingungen für die Wählbarkeit von Zubehör basieren auf Produkteigenschaften oder Anlagenparametern [1].

3.3 Beispiel: Wärmepumpen (VDI 3805 Blatt 22)

Blatt 22 konkretisiert dies für Wärmepumpen [2] durch spezifische Datenkategorien: Betriebsdaten umfassen Korrekturfaktoren, Temperaturdifferenzen und Füllmenge des Kältemittels. Leistungsdaten beschreiben Heiz- und Kühlleistung, elektrische Leistungsaufnahme, Leistungszahl sowie klimabedingte ErP-Daten. Physikalische Eigenschaften betreffen Schallleistungspegel, Schutzart (IP-Code) und Masse. Anschlüsse werden durch spezifische Bezeichnungen wie ELU (Eintritt Luft) oder HV (Heizungs-Vorlauf) charakterisiert. Steuerungs- und Regelungsaspekte erfassen Anzahl der Betriebsprogramme, Leistungsstufen sowie Integration von Warmwasser-, Solar-, Mischer- oder Kühlregelungen einschließlich BUS-Fähigkeit. Dimensionale Angaben umfassen spezielle Einbring-, Aufstell- oder Kippmaße. Gaswärmepumpen-spezifische Daten beinhalten Gasart, Nennwärmebelastung und Gas-Anschlusswert. Die Betriebsweise wird differenziert nach monovalent, monoenergetisch, bivalent-alternativ, bivalent-parallel oder bivalent-teilparallel.

4 Implementierung

Dieser Abschnitt beschreibt Modellierung (Pydantic) und Visualisierung (Dash Cytoscape) anhand eines Wärmepumpen-Datensatzes.

4.1 Datenmodellierung und Validierung

Jede VDI-Satzart wird als Pydantic-Klasse implementiert (Algorithm 1). 54 Satzarten wurden modelliert (5 Basisklassen, 33 allgemeine VDI 3805-Klassen aus der Grundlagennorm und 16 wärmepumpenspezifische Klassen aus VDI 3805-22). Die Implementierung validiert automatisch Typenkonformität beim Einlesen der VDI-Datensätze. Für die strukturierte Modellierung und formale Beschreibung der Produktdaten wird in dieser Arbeit *Pydantic* verwendet. Obwohl *Pydantic* selbst nicht in den vorliegenden Normen referenziert wird, bietet es eine robuste Grundlage, um die komplexen Datenstrukturen der VDI 3805-Richtlinien in maschinenlesbarer und validierbarer Form abzubilden. *Pydantic* ermöglicht die Definition von Datenmodellen mit Typ-Annotationen, wodurch eine automatische Datenvalidierung beim Einlesen der VDI-Datensätze gewährleistet wird. Dies ist entscheidend, um die „syntaktische Richtigkeit und semantische Sinnfälligkeit“ der Daten zu prüfen, wie sie auch für VDI 3805-Dateien gefordert ist. Jede Satzart (010 Vorlauf, 100ff. Klassifikation, 700 Merkmals-/Produktelementdaten, 970 Geometrie) wird als eigene *Pydantic*-Klasse abgebildet (Algorithm 1). Besondere Bedeutung kommt dabei der Abbildung der Hierarchie der Produktstruktur zu, die in VDI 3805 durch die TGA-Nummer exakt wiedergegeben wird. *Pydantic-Modelle* können diese hierarchischen Beziehungen durch geschachtelte oder referenzierende Datenstrukturen abbilden, was die Einhaltung der VDI-Modellierungslogik unterstützt.

4.2 Interaktive Graphvisualisierung

Für die explorative Darstellung der zugrunde liegenden Beziehungen der Produktdaten wurde ein interaktiver Graph-Viewer auf Basis von *Dash Cytoscape* implementiert. Die Notwendigkeit einer visuellen Darstellung ist im Kontext des Produktdatenaustauschs hoch, da herkömmliche Kataloge Daten in Tabellen, Diagrammen und Zeichnungen veranschaulichen. Die komplexen hierarchischen Abhängigkeiten und Modellierungsprinzipien in VDI-Daten erfordern eine transparente und zugängliche

Algorithm 1: Beispiel einer Pydantic-Klasse

```
class MainProductGroup(RecordWithIndex):
    record_type:
        Literal['100', '110'] = Field(alias="1")
    sort_number_for_display_sequence:
        Optional[int] = Field(alias="3", default=None)
    product_name:
        Optional[str] = Field(alias="4", default=None)
    main_product_group:
        Optional[str] = Field(alias="5", default=None)
    product_classification:
        Optional[str] = Field(alias="6", default=None)
    creation_date:
        Annotated[Optional[date], BeforeValidator(a8_to_date), Field(alias="7",
        default=None)]
    comment:
        Optional[str] = Field(alias="8", default=None)
    index_of_the_start_data_record_960_media_data:
        Optional[int] = Field(alias="9", default=None)
```

Darstellung. Der Graph-Viewer bildet die unterschiedlichen Satzarten und ihre Beziehungen als Knoten und Kanten ab. Dies ermöglicht es, die „hierarchischen Abhängigkeiten“ innerhalb der VDI-Daten „visuell darzustellen“. Beispielsweise werden die Produkthauptgruppen und Produktvarianten (Satzarten 100 bis 560) als hierarchisch angeordnete Knoten visualisiert, was die Struktur des Produktdatenmodells gemäß VDI 3805 Blatt 1 nachvollziehbar macht. *Dash Cytoscape* bietet die Möglichkeit, folgende Aspekte interaktiv zu untersuchen:

- **Hierarchische Gliederungen:** Die „logisch entwickelte Nummer, die exakt die Hierarchie der Produktstruktur wiedergibt“ (TGA-Nummer) kann durch die Anordnung von Knoten und Kanten dargestellt werden.
- **Beziehungen zwischen Informationseinheiten:** Die VDI 3805 Blatt 1 definiert Beziehungen wie „Eine Informationseinheit A hat eine Beziehung zu einer oder mehreren Informationseinheiten B“, die im Graphen als gerichtete Kanten visualisiert werden.
- **Abhängigkeiten von Zubehören:** Die in Satzarten 930 bis 930.04 beschriebenen, beliebig tief gestaffelten Abhängigkeitsbäume von Zubehören können als verschachtelte oder verbundene Knotengruppen dargestellt werden.
- **Verweise auf externe Dateien:** Mediendaten (z. B. Fotos, Videos, Tondokumente), die in VDI-Dateien lediglich referenziert und nicht integriert werden, können als Links visualisiert werden, die auf externe Ressourcen zeigen.

4.3 Exemplarische Anwendung: VDI 3805 Wärmepumpen-Datensatz

Als exemplarische Anwendung dient ein Wärmepumpen-Datensatz. Die Produktstruktur ist hierarchisch (siehe Abbildung 1). Im Viewer werden folgende Elemente abgebildet:

- **Produkthauptgruppen und Varianten:** Die Satzarten 100 (Einsatzbereich), 110 (Typ), 200 (Wärmequelle), 500 (Antriebsenergie) und 550 (Funktionsprinzip) definieren die Hauptklassifikationen einer Wärmepumpe. Diese werden als übergeordnete Knoten dargestellt, unter denen sich die spezifischen Werte (z. B. „Heizung und Trinkwassererwärmung“ für Einsatzbereich, „Luft-Wasser“ für Typ, „Außenluft“ für Wärmequelle oder „elektrischer Strom“ für Antriebsenergie) als untergeordnete Knoten oder Merkmale finden.

- **Technische Merkmale:** Produkte werden durch Merkmalswerte beschrieben. Die Satzarten 700 ff. in VDI 3805 Blatt 22 enthalten detaillierte technische Daten wie Heizleistung, Leistungszahl, elektrische Leistungsaufnahme, Kältemitteldetails, ErP-Daten, Einbringmaße und Steuerungsinformationen. Diese werden als Eigenschaften der Produktknoten oder als eigene Untergraphen dargestellt.

    · **Statische Merkmale:** Repräsentieren konstante Charakteristiken (z. B. Abmessungen wie Länge, Breite, Höhe). Diese Werte sind direkt im Katalog hinterlegt.

    · **Dynamische Merkmale:** Beschreiben das Verhalten des Produkts unter Betriebsbedingungen (z. B. Druckverlust, Schallleistungspegel). Ihre Werte sind nicht fest im Katalog zugeordnet, sondern werden durch Implementierung einer Funktion berechnet, die von bestimmten Parametern abhängt.

- **Funktionen und Algorithmen:** In VDI 3805 Blatt 1 sind Satzarten 600 (Deklarationen) und 610 (Definitionen) vorgesehen, um „Berechnungsformeln und Algorithmen zur Auslegung und Berechnung des Produkts“ zu übermitteln. Im Graphen können diese Funktionen als spezielle Knoten dargestellt werden, die auf die zugehörigen technischen Merkmale verweisen, welche als Parameter dienen. Die Satzart 820 in VDI 3805 Blatt 1 referenziert Funktionen, die virtuelle Datensätze der Satzarten 800 (TGA-Nummer) und 810 (Artikelnummern) generieren, um die Datenmenge zu reduzieren. Diese Generierungslogik kann ebenfalls als Beziehung im Graphen visualisiert werden.

- **Geometriedaten:** VDI 3805 Blatt 1 beschreibt in Satzarten 970 ff. die geometrische Gestalt einer TGA-Baueinheit, einschließlich Störräumen, Anschlussdaten (970.01), Gestaltdaten (970.02, 970.03), Materialdaten (970.04) und Symboldaten (970.05, 970.06). Besonders die Anschlüsse sind für die automatische Installation und geometrische Prüfung im CAD-System relevant. Sie werden mit einer ID, Funktion, Position, Richtung und Maßen beschrieben. Im Graph-Viewer können diese Geometriedaten als spezifische Unterknoten der Produktknoten dargestellt werden, was die detaillierte geometrische und funktionale Konnektivität des Produkts veranschaulicht.

Dies unterstützt das spätere Mapping Richtung DIN EN ISO 16757.

4.4 Einordnung gegenüber bestehenden Werkzeugen

Es existieren verschiedene Ansätze zur Nutzung technischer Produktdaten:

- **Proprietäre Katalogbrowser** / **Hersteller-CDs**: Fokus auf tabellarische Listen, eingeschränkte semantische Navigation; meist keine explizite Abbildung von Abhängigkeitsbäumen.

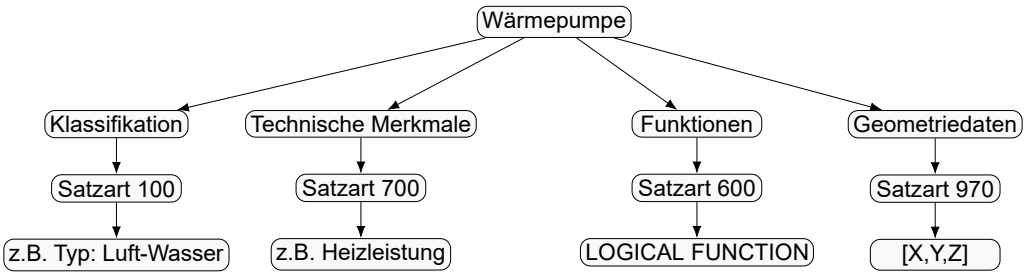


Abbildung 1: Vereinfachte Darstellung der VDI 3805 Wärmepumpen-Datenstruktur mit exemplarischen Satzarten.

- **BIM-/CAD-Familien (z.B. generische Revit-Familien)**: Geometrie + wenige Parameter; Variantenexplosion statt regelbasierter Reduktion; semantische Beziehungen implizit.
- **Standard-Datenwörterbücher (z.B. bSDD)**: Begriffliche Harmonisierung; keine interaktive Visualisierung konkreter Kataloginstanzen.
- **Graphdatenbank-Demos**: Zeigen Netzwerke, jedoch ohne domänenspezifische Satzart-Semantik von VDI 3805.

Der vorliegende Ansatz kombiniert (i) formale Satzart-Validierung, (ii) explizite Hierarchie- und Funktionsknoten und (iii) dynamische Zubehörabhängigkeiten in einer integrierten Graphsicht. Damit adressiert er eine Lücke zwischen statischer Tabellen-Navigation und generischer Graphvisualisierung.

5 Ergebnisse

Die Anwendung auf einen konkreten Wärmepumpen-Datensatz mit über 23.000 Zeilen demonstrierte die erfolgreiche vollständige Validierung aller VDI-Datensätze mittels der 54 implementierten Pydantic-Satzarten. Der generierte Graph umfasste entsprechend über 23.000 Knoten, wobei jede validierte Zeile als separater Knoten dargestellt wurde. Die aktuelle Implementierung zeigt alle Datensätze als isolierte Knoten ohne Kanten-Verbindungen, da die automatische Ableitung der semantischen Beziehungen zwischen VDI-Satzarten noch in Entwicklung ist. Die Navigation erwies sich aufgrund der schieren Anzahl gruppierter Knoten als herausfordernd, was die Notwendigkeit fortgeschrittener Layout-Algorithmen und Filterungsmechanismen verdeutlicht.

6 Diskussion und Ausblick

Die Validierung von über 23.000 VDI-Datensätzen demonstriert die Praktikabilität des Pydantic-basierten Ansatzes für große Datenmengen. Jedoch zeigt die Erfahrung mit der Graphvisualisierung, dass die reine Abbildung aller Datensätze als isolierte Knoten ohne semantische Kanten-Verbindungen die Navigierbarkeit stark einschränkt. Diese Erkenntnis unterstreicht die zentrale Herausforderung: die automatische Ableitung hierarchischer und semantischer Beziehungen aus VDI-Strukturen.

Nächste Schritte: Implementierung von Kanten-Generierung basierend auf Hierarchien durch die TGA-Nummern und Satzart-Referenzen, fortgeschrittene Layout-Algorithmen (Clustering, Lazy Loading), halbautomatisches Mapping zu DIN EN ISO 16757. **Erkannte Limitationen:** Skalierungsproble-

me bei großen Graphen, fehlende semantische Kanten-Ableitung, Navigation in dichten Knotengruppen.

7 Fazit

Das Konzept zeigt die erfolgreiche formale Validierung großer VDI 3805-Datensätze und identifiziert gleichzeitig kritische Herausforderungen bei der Visualisierung komplexer Produktdatenstrukturen. Die Arbeit leistet einen wichtigen Beitrag zur Strukturanalyse von VDI-Daten und schafft eine solide Grundlage für die Entwicklung verbesserter Visualisierungs- und Transformationsansätze. Zukünftige Arbeiten fokussieren auf die Lösung der Skalierungs- und Navigationsherausforderungen sowie die automatisierte Ableitung semantischer Beziehungen für das Mapping zu DIN EN ISO 16757.

Acknowledgements

Diese Arbeit entstand im Rahmen des Projekts "Parametrisches BIM für die Wärmewende", gefördert durch das BBSR im Programm Zukunft Bau. Besonderer Dank gilt den Projektpartnern für die bereitgestellten Produktdaten sowie Malte Junge, Noemi Kremer und Prof. Jakob Beetz für ihre Unterstützung.

Literatur

[1] *Produktdatenaustausch in der technischen Gebäudeausrüstung - Grundlagen*, Düsseldorf: Verein Deutscher Ingenieure (VDI), 2022.

[2] *Produktdatenaustausch in der technischen Gebäudeausrüstung - Wärmepumpen*, Düsseldorf: Verein Deutscher Ingenieure (VDI), 2019.

[3] *Datenstrukturen für elektronische Produktkataloge der Technischen Gebäudeausrüstung - Teil 1: Konzepte, Architektur und Modelle*, Berlin: Deutsches Institut für Normung (DIN), Okt. 2019.

[4] *Datenstrukturen für elektronische Produktkataloge der Technischen Gebäudeausrüstung - Teil 2: Geometrie*, Berlin: Deutsches Institut für Normung (DIN), Feb. 2019.

[5] *Datenstrukturen für elektronische Produktkataloge der Technischen Gebäudeausrüstung - Teil 4: Datenwörterbücher für Produktkataloge*, Berlin: Deutsches Institut für Normung (DIN), Apr. 2024.

[6] *Datenstrukturen für elektronische Produktkataloge der Technischen Gebäudeausrüstung - Teil 5: Austauschformat für Produktkataloge*, Berlin: Deutsches Institut für Normung (DIN), Mai 2024.

[7] N. Kremer, J. Beetz und E. Hjelseth, »Identification and exploration of data resources in product configuration processes«, 2024.

Construction site obstacle detector assisted by large language models

Miguel Guzmán-Merino[1] and Dennis Grauer[1]
[1]AI for Sustainable Construction, University of Rostock
E-mail(s): miguel.merino@uni-rostock.de, dennis.grauer@uni-rostock.de

Abstract: Construction sites present situations where obstacles interrupt the desired task. For some robot application, these obstacles cannot simply be avoided. In those cases, it is necessary to remove the obstacles to perform the desired task. Information like the spatial context, the weight and dimensions of the obstacles, or even the specifications of the robots is needed if the manipulation of the obstacle is intended to be automated. In this paper, we propose a system for the identification of obstacles and their attributes in a desired workspace. The user provides the workspace specifications and a monocular RGB image as inputs to the system. Using Depth Anything and Open3D, we recreate and analyse the environment captured in the monocular RGB image. To extend the information that cannot be derived from the RGB image, we proposed the use of a Large Language Model (LLM) assistant. The goal of our system is to provide information about the obstacles in a defined workspace.

Keywords: Depth estimation, Point cloud segmentation, Large language models, Construction robotics

DOI: 10.18154/RWTH-CONV-254891. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Different implementations of robotic systems in construction have already been developed. Tasks like wall painting or sanding have been successfully automated, and different products can be found on the market. Several developments in autonomous bricklaying, concrete printing, or timber construction can also be found in the research literature. However, these systems still require human intervention at different stages of task execution, for instance: (i) configuration of the task parameters, (ii) preparation of the workspace, or (iii) the setup of the robots in the workspace. Therefore, it is necessary to target these aspects to increase the automation of robotic systems on construction sites. In this paper, we focus on the preparation of the workspace in indoor environments, a topic that has already been studied during the early stages of outdoor construction sites [1]. More precisely, we focus on the identification of obstacles and their relevant attributes.

Since preparing a workspace involves manipulating obstacles, it is necessary to establish the relevant information for this manipulation. This information can be derived from related work such as [2]. In this study, the authors present a system for robotic manipulation tasks in environments without references. For object manipulation, the authors provide the robot with information such as the location,

dimensions, and weight of the objects. This information is usually available or can be accurately estimated in controlled environments. However, this information is difficult to obtain in construction environments where little or no prior information is available and the obstacles that will be encountered cannot be anticipated. To obtain the necessary information, we identify three aspects that need to be addressed: (i) the delimitation of the work area to differentiate between objects in the environment and obstacles, (ii) the identification of the objects in the workspace, and (iii) the estimation of the parameters of the obstacles.

To address these aspects, we propose a system that uses a single monocular RGB image as a source of information about the real environment, together with a user request containing details about the workspace. These initial constraints allow us to minimize human intervention in the system workflow. In this use case, a user does not need training to operate complex and expensive equipment such as LiDAR or depth cameras; a single image captured on a phone will suffice to derive the necessary information. We explore depth estimation and point cloud segmentation techniques to reconstruct and analyse the scenario, in combination with large language model (LLM) assistants for obstacle identification and the enrichment of information about them.

2 Related work

Extensive literature exists on image processing systems, providing solutions for object identification and pose estimation. Systems like the one proposed by [3] use markers to define reference points in images. Although this approach offers high accuracy and reliability, it cannot be directly applied in real construction sites due to the inability to ensure marker visibility. Moreover, this system requires human intervention to place the markers beforehand.

Image object segmentation and identification using machine learning algorithms also provide a method for extracting information without relying on references. Tools like You Only Look Once (YOLO) offer well-supported solutions for multiple image processing problems, even in the construction domain [4]. However, these models require prior knowledge about the objects to be identified. This is not feasible when dealing with obstacles in construction sites because obstacles cannot always be anticipated. Furthermore, 2D image analysis does not provide the spatial information necessary for object manipulation.

Approaches such as [5], [6] that do not rely on prior knowledge of the objects in the scene can also be found. In [5], the authors present a point cloud segmentation technique guided by natural language for semantic scene descriptions of point clouds. In [6], a multi-modal large language model (LLM) for point cloud understanding is used, including open vocabulary object recognition, scene captioning, and question answering about measured 3D data. Although these approaches achieve high performance for scene description tasks, some points differ from our approach. (i) These approaches use high-quality point cloud data, primarily sourced from multi-sensor systems or comprehensive synthetic datasets. In contrast, our approach is limited to point cloud reconstruction from a single monocular RGB image of a real scene. (ii) The identification in [5] requires natural language guidance, whereas our approach does not rely on such input. (iii) Our approach is not focused on scene description but rather on the identification of obstacles and their attributes.

3 Proposed system

3.1 System overview

To describe the workflow, we propose a use case scenario. A user located at a construction site takes a picture of the scene and provides a request to the system. In the picture, the target element—such as a wall—is captured, and the request "Paint 1 m by 1.5 m area of the wall 1 m from the right corner" contains the information about the task and the necessary workspace. With this information, the system identifies the objects in the corresponding workspace and their attributes. The result information could be provided to a robotic system that would subsequently remove the obstacles. Figure 1 shows the proposed system workflow. As shown in the system diagram, a list of available tools for obstacle manipulation is also provided to the system. In the following sections, the modules integrating the system are described in detail, along with a use case example.

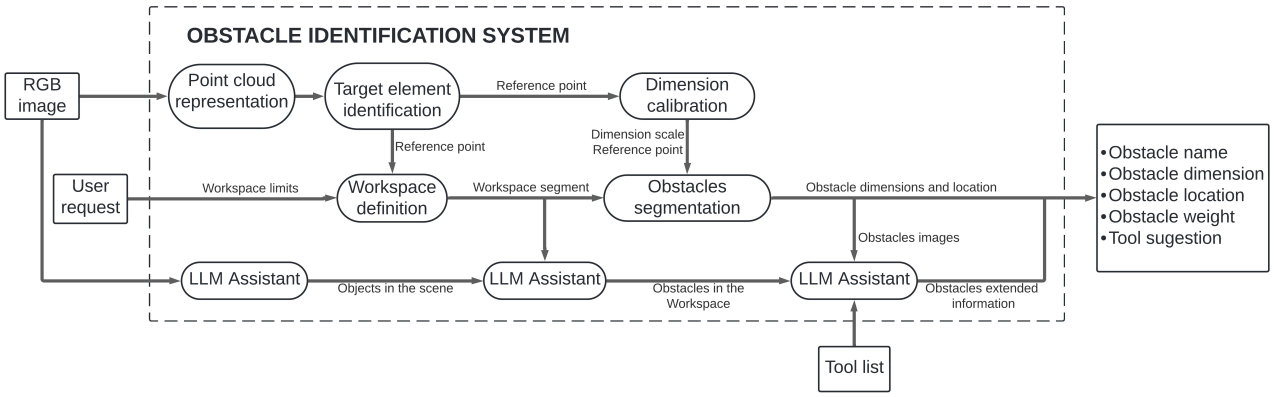


Figure 1: Workflow diagram of the system. The diagram shows how the external information of the system is treated in the different modules to obtain the desired information about the obstacles.

3.2 Point cloud representation

Taking the RGB image as input, the point cloud representation module reconstructs the real scenario in a 3D format. This representation is subsequently used for the identification of the target element and the obstacles, and the definition of the working area. To create the 3D point cloud representation from the RGB image, two steps were performed: the estimation of depth information and the 3D point cloud reconstruction. To obtain the depth information, we used Depth Anything [7], a foundation model trained on labelled images that generates depth maps from monocular images. Figure 2 (middle image) represents the depth estimation of the RGB image shown on the left. To create the point cloud reconstruction, we used Open3D scripts that relate the pixel information of the RGB image with the depth information from the depth map estimation. Alternative point cloud reconstruction techniques using multiple images provide more accurate estimations of real measurements through feature recognition approaches [8], [9]. However, these approaches require several images or video collection, increasing user participation and computational power. Instead, we aim to study the case where the user needs to provide only a single image.

3.3 Target element identification

In this module, the identification of the target element and the referent points is performed. We define the target element as the element where the final task will be executed. We assume that, in

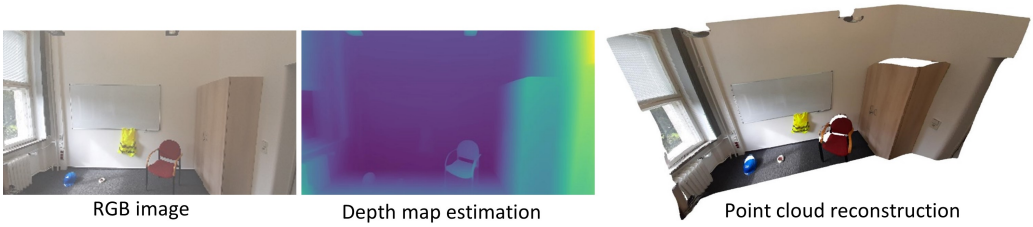


Figure 2: Reconstruction of the point cloud representation of the real scenario. On the left, the proposed RGB monocular image represents proposed scenario . In the middle, the depth map reconstruction of the scene. On the right, the point cloud reconstruction of the scenario

construction environments, possible target elements are structural components such as walls, floors, or ceilings. These elements can be identified as planar surfaces in the point cloud reconstruction.

To identify the target element in the scene, it must be captured entirely in the source image. As the target element is the most relevant planar surface in the scene, it can be identified using the plane segmentation tool from Open3D. As shown in Figure 3, we identify the target element in the point cloud representation (left) and isolate it along with its bounding box (right). The bounding box contains the *reference point* that is used to locate objects in the scene. In this case, we establish the *reference point* at the lower right corner of the target element. Different point cloud segmentation approaches, such as those described in [10], [11], could also be applied. However, we opt for the planar segmentation function included in Open3D because it provides a simple and direct solution to our problem.

Figure 3: Target element detection. Planar segmentation is applied to detect the corresponding target element, the highlighted wall in the proposed scenario on the left side. On the right side, the isolated target element with the corresponding bounding box.

3.4 Dimension information

Since no real dimensions can be extracted from the provided source image, we perform a calibration process using a known scenario. For this calibration, we locate different objects in the scene and record their real positions relative to the *reference point*, defined as the lower right corner of the target element. By identifying this *reference point* in the point cloud reconstruction, it is possible to determine the appropriate *scale factor* that relates the real positions to the corresponding locations of the objects in the point cloud representation.

3.5 Workspace definition

We define the workspace as the necessary space around the target element to execute the desired task. To identify the workspace in the point cloud reconstruction, this module uses the *reference point*, the *scale factor*, and the information in the user request related to the workspace. For instance, given

Figure 4: Workspace definition and obstacles identification. The necessary workspace defined in the user request is represented in the point cloud reconstruction on the left side image and isolated in the middle image. On the right side, each of the obstacles in the workspace is isolated.

the user request "Paint 1 m by 1.5 m area of the wall 1 m from the right corner," a workspace of 1 meter by 1.5 meters needs to be defined at the corresponding location in the point cloud representation. Using the *scale factor* and the *reference point*, we define the desired workspace in the point cloud representation using Open3D scripts, as shown in Figure 4. The left side of the figure shows the workspace in the point cloud reconstruction for the proposed request, the middle shows the isolated working area, and the right displays the section of the point cloud further analyzed to identify obstacles within the working area

3.6 Obstacle segmentation

To identify the obstacles in the defined working space, a new segmentation process is applied to the point cloud section of the workspace (Figure 4, middle image). By applying plane segmentation techniques, it is possible to remove the remaining planar surfaces (wall and floor). The remaining points represent the obstacles in the workspace. The result of the segmentation process is shown in Figure 4, right side. To extract each object individually, we use the clustering tool included in Open3D. This process allows us to isolate each obstacle and compute their corresponding bounding boxes, which contain information about the location and dimensions of the objects. We refer the location of the obstacles to the *reference point*. Afterwards, we estimate the location and dimensions in the real world by applying the *scale factor* to the locations and dimensions in the point cloud representation.

3.7 LLM assistant

In this module, the information about the location and dimensions of the obstacles is extended to include their identification as well as an estimation of their weight, parameters that can not be derived from the analysis of the scenario. As part of this information extension, a suggestion of the most suitable tool in the robotic system for manipulating each object is also provided. For obstacle identification and information extension, we propose using an LLM assistant. We prompt the assistant with the available information at different steps of the process to guide the identification of obstacles.

As shown in the system diagram in Figure 1, there are 3 calls to the LLM assistant: (i) The assistant is requested to identify the objects in the source image. With this initial evaluation, the assistant creates a list of objects that is stored as historical information. (ii) An image of the workspace, represented in Figure 4, is provided to the assistant along with the list of objects identified in the first call. In this case, the assistant is asked to identify the objects in the workspace, based on the information stored from the previous request. Thanks to this historical information, the LLM assistant can properly identify objects even if they have been partially cut during the segmentation of the working area. This information is also stored. (iii) The LLM assistant is provided with an image of each obstacle obtained

Request 1:	Request 2:	Request 3:
Provide a list of the objects in the picture. Assistant response: Here's a list of the main objects visible in the room: • Wooden storage cabinet • Wall-mounted whiteboard • Yellow high-visibility safety vest (hanging on the whiteboard) • Red upholstered chair with wooden armrests • Roll of adhesive tape (on the floor) • Blue safety helmet / hard hat (on the floor) • Ceiling-mounted fluorescent light fixtures • Window with blinds • Radiator (under the window) • Electrical outlets and power sockets (on the wall)	Which of the objects are in the new picture? Assistant response: The following objects from the original list appear in the new picture: • Yellow high-visibility safety vest • Red upholstered chair (arm and seat visible) • Roll of adhesive tape (on the floor) • Blue safety helmet / hard hat • Dark carpeted floor	Which object is this one? Provide the name of the object. Provide an estimation of the weight of the object and the best tool from the Tool list to manipulate it. Assistant response: High-visibility safety vest. 0.3 Kg. Mechanical Gripper Blue safety helmet. 0.4 Kg. Vacuum Gripper Roll of adhesive tape. 0.1 Kg. Mechanical Gripper. Red upholstered chair. 5.0 Kg. Mechanical Gripper

Figure 5: LLM assistant requests and responses. Each of the boxes displays the request and response in each of the calls done to the LLM assistant.

during obstacle segmentation (Figure 4, right side) and is required to identify the object according to the information stored from the second request. After identifying an obstacle, the assistant is also asked to provide an estimation of its weight based on the corresponding dimensions, as well as to recommend the best tool from the list of available tools to manipulate it.

An example of the sequence of requests and responses to the assistant can be shown in Figure 5. The output of this module is the output of the system. It is all the information related to each of the identified obstacles: obstacle name, dimensions, weight estimation, locations and a suggestion of the best tool to manipulate it. We propose the output of the system as the input of a task allocation system where the appropriate robot manipulator is selected and the information about each of the obstacles is used to remove them from the workspace.

During our experiments, we did not observe hallucinations like incorrectly identified objects or tools. Our proposed process chain bears the risk of error propagation of those incorrect labels. Further development will be necessary in this aspect.

4 Results

As discussed in Section 3.7, the LLM assistant accurately identifies obstacles thanks to the historical knowledge gained during the different stages of the identification process. To evaluate the accuracy of the estimated locations and dimensions, five experiments were conducted. A total of five objects were used in the experiments, placed at different locations within the scenario presented in Figure 2. The requested workspace was also modified between experiments. Table 1 summarizes the results of the estimated locations and dimensions.

The locations and dimensions presented in Table 1 are estimated from the corresponding bounding box of each object. The highlighted experiments correspond to misclassification. In experiment 1, object 3 is partially represented in the selected workspace, and the bounding box does not contain the entire object. In experiment 4, objects 1 and 5 are detected as a single object. These results highlight a key limitation of the proposed system. Although the segmentation process is generally sufficient to isolate each obstacle, a lack of precision can cause misclassification problems. The bounding boxes

Table 1: The table presents the results of the estimated location and dimensions of the detected obstacles in the scene.

	Obstacle	Estimate Location [m] (x,y,z)	Real Location [m] (x,y,z)	Location error [abs] (x,y,z)	Estimate dimension [m] (h,d,h)	Real dimension [m] (h,d,h)	Dimension error [abs] (h,d,h)
Experiment 1	object 1	(1.61,0.08,0.17)	(1.55,0.07,0.12)	(0.064,0.014,0.046)	(0.57,0.14,0.26)	(0.25,0.13,0.21)	(0.320,0.010,0.050)
	object 2	(0.96,0.06,0.47)	(0.95,0.07,0.4)	(0.006,0.014,0.065)	(0.17,0.17,0.29)	(0.20,0.16,0.30)	(0.030,0.010,0.010)
	object 3	(1.34,0.96,0.10)	(1.20,1.30,0.10)	(0.140,0.340,0.000)	(0.32,0.15,0.24)	(0.50,0.65,0.15)	(0.180,0.500,0.090)
Experiment 2	object 4	(1.53,0.17,0.53)	(1.70,0.70,0.35)	(0.170,0.530,0.180)	(1.38,0.26,0.84)	(0.60,0.15,0.80)	(0.780,0.110,0.040)
	object 5	(0.87,0.18,0.7)	(0.90,0.16,0.50)	(0.030,0.020,0.200)	(0.19,0.15,0.34)	(0.18,0.16,0.25)	(0.010,0.010,0.090)
Experiment 3	object 1	(1.02,0.07,0.36)	(1.00,0.07,0.40)	(0.020,0.000,0.040)	(0.52,0.09,0.48)	(0.25,0.13,0.21)	(0.270,0.040,0.270)
	object 5	(1.77,0.07,0.34)	(1.90,0.07,0.35)	(0.130,0.001,0.015)	(0.17,0.12,0.26)	(0.18,0.16,0.25)	(0.010,0.040,0.010)
Experiment 4	object 4	(0.82,0.24,0.16)	(1.10,0.30,0.10)	(0.280,0.060,0.060)	(0.60,0.50,0.18)	(0.8 ,0.60,0.15)	(0.200,0.100,0.030)
	object 1	(1.70,0.03,0.38)	(2.30,0.07,0.30)	(0.600,0.040,0.080)	(0.47,0.14,0.24)	(0.25,0.13,0.21)	(0.180,0.160,0.250)
	object 5		(2.00,0.07,0.30)	(0.300,0.040,0.080)		(0.18,0.16,0.25)	(0.280,0.060,0.020)
Experiment 5	object 4	(1.37,0.15,0.54)	(1.70,0.07,0.35)	(0.330,0.080,0.190)	(0.88,0.21,0.82)	(0.60,0.15,0.80)	(0.280,0.060,0.020)
	object 1	(2.00,0.09,0.54)	(2.50,0.07,0.25)	(0.500,0.020,0.290)	(0.40,0.15,0.43)	(0.25,0.13,0.21)	(0.150,0.020,0.220)
	object 5	(0.71,0.16,0.69)	(0.90,0.07,0.50)	(0.190,0.090,0.190)	(0.18,0.14,0.34)	(0.18,0.16,0.25)	(0.000,0.020,0.090)

created around the obstacles may include points that do not belong to the objects, leading to incorrect estimations of the obstacles' locations and dimensions.

5 Conclusion

In this paper, a system for obstacle identification has been presented. The system primarily focuses on low user requirements. Using a single monocular image as the sole source of information about the real-world scenario, we apply depth estimation techniques to reconstruct the environment in a 3D point cloud representation. Combining the point cloud reconstruction with the information in the user request, we identify obstacles in the desired workspace and their attributes of interest. This approach enables non-expert users to interact with complex robotic systems, as the extracted information could be used as input for a robotic system to manipulate the identified obstacles. With this proposal, we also address the heterogeneity of construction environments. The object recognition performed by the LLM assistant allows the identification of a large number of obstacles without anticipating or providing information about them. This flexible object identification would not be possible using object identification systems trained on a predefined set of images. The LLM assistant also provides information that cannot be derived from scenario analysis.

Although the proposed system presents promising results, further developments are necessary. As discussed in Section 4, a more precise segmentation process is mandatory. In future work, we will explore alternative techniques for point cloud segmentation, such as colour-based segmentation. We will also investigate the inclusion of entire objects as obstacles, even when they are only partially included in the defined workspace. We will also further develop the identification process to address hallucinations and error propagation. We will include a list of expected objects or a feasibility-check step; either can be performed by an LLM.

References

[1] O. Elmakis, T. Shaked, and A. Degani, "Vision-based uav-ugv collaboration for autonomous construction site preparation", *IEEE Access*, vol. 10, no. 1, pp. 51 209–51 220, 2022. DOI: 10.1109/ACCESS.2022.3170408

[2] A. Collet, D. Berenson, S. S. Srinivasa, and D. Ferguson, “Object recognition and full pose registration from a single image for robotic manipulation”, in *2009 IEEE International Conference on Robotics and Automation*, 2009, pp. 48–55. DOI: 10.1109/ROBOT.2009.5152739

[3] C. Feng, Y. Xiao, A. Willette, W. McGee, and V. R. Kamat, “Vision guided autonomous robotic assembly and as-built scanning on unstructured construction sites”, *Automation in Construction*, vol. 59, pp. 128–138, 2015. DOI: https://doi.org/10.1016/j.autcon.2015.06.002

[4] F. Hao, T. Zhang, G. He, R. Dou, and C. Meng, “Casnli-yolo: Construction site multi-target detection method based on improved yolov5s”, *Measurement Science and Technology*, vol. 35, no. 8, 2024.

[5] S. He, H. Ding, X. Jiang, and B. Wen, “Segpoint: Segment any point cloud via large language model”, in *European Conference on Computer Vision*, Springer, 2024, pp. 349–367.

[6] R. Xu et al., “Pointllm-v2: Empowering large language models to better understand point clouds”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025.

[7] L. Yang, B. Kang, Z. Huang, X. Xu, J. Feng, and H. Zhao, *Depth anything: Unleashing the power of large-scale unlabeled data*, 2024. arXiv: 2401.10891. [Online]. Available: https://arxiv.org/abs/2401.10891

[8] M. Brown and D. Lowe, “Unsupervised 3d object recognition and reconstruction in unordered datasets”, in *Fifth International Conference on 3-D Digital Imaging and Modeling (3DIM’05)*, 2005, pp. 56–63. DOI: 10.1109/3DIM.2005.81

[9] X. Zhang, J. T. Blocksom, and D. D. Miller, “Large-scale urban environment modeling from videos using image content segmentation and alignment”, in *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops*, 2009, pp. 1848–1854. DOI: 10.1109/ICCVW.2009.5457507

[10] F. Engelmann, T. Kontogianni, J. Schult, and B. Leibe, “Know what your neighbors do: 3d semantic segmentation of point clouds”, in *Computer Vision – ECCV 2018 Workshops*, L. Leal-Taixé and S. Roth, editors, Cham: Springer International Publishing, 2019, pp. 395–409.

[11] F. Verdoja, D. Thomas, and A. Sugimoto, “Fast 3d point cloud segmentation using supervoxels with geometry and color for 3d scene understanding”, in *2017 IEEE International Conference on Multimedia and Expo (ICME)*, 2017, pp. 1285–1290. DOI: 10.1109/ICME.2017.8019382

Short-term HVAC load prediction using Temporal Graph Convolutional Networks

Dennis Grauer[1], Miguel Guzman-Merino[1] and Nils Soeren Krause[1]
[1]Artificial Intelligence for sustainable construction, University of Rostock, Germany
E-mail(s): dennis.grauer@uni-rostock.de, miguel.merino@uni-rostock.de, nils.krause@uni-rostock.de

Abstract: Machine learning approaches for building performance prediction have gained popularity in recent years. Backed by the higher availability of Internet of Things sensor data, researchers have created well-performing data-driven prediction models for indoor climate conditions, energy usage and thermal load. These models use features highly correlated to their prediction target to score high accuracies, but this often leads them to overfit, struggling with unseen data and not producing outputs complying with physical laws. In this paper, we present a Temporal Graph Convolutional Network (TGCN) approach, seeking to overcome these limitations. The model is intended to predict short-term Heating, Ventilation and Cooling (HVAC) load for a office building. Instead of using few features, we incorporate a high number of sensors into the models underlying graph structure. We validate this approach on a three year data set of an office building. The data set contains data from the 2020 Covid-19 pandemic, where occupational patterns of the building change and the correlations governing the HVAC load shifts. The TGCN model is trained with pre-pandemic data and confronted with the pandemic situation. The results show, that the model is able to maintain its performance on the prediction task for out-of-distribution data.

Keywords: Graph neural networks, building energy, machine learning

DOI: 10.18154/RWTH-CONV-254882. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Although predictions about building performance traditionally relied on physical simulations, the wider spread of Internet of Things (IoT) sensors has encouraged researchers to develop data driven prediction models. While these data driven approaches offer a high accuracy and require less configuration and knowledge about the buildings characteristics, they fall short of physics-based models in two categories: (i) as these models are purely based on correlation, they tend to overfit on correlation and struggle with capturing causation, particular in unseen situations, (ii) they are considered black-box models, as the interpretability of their predictions falls short to physic-based models [1]. In this paper, we want to focus on the first problem.

Neural networks conducting their learning on graph structures, commonly referred to as Graph Neural Networks (GNNs), have received extensive attention in the field of building (energy) management.

When their graphs encode physical relations, they can be categorized as a kind of physics-informed machine learning.

We want to explore the potential of graph neural networks to deal with a large, physics based graph structure. Therefore, we build a Temporal Graph Convolutional Network (TGCN) for a real-world office building dataset, aiming to predict short-term HVAC load.

The graph is based on the physical relations governing the HVAC load and utilizes a high number of IoT sensors to make its predictions. We seek to find an approach, that performs well on unseen data by learning causation instead of correlation. In the following section, we will first look at the relevant literature. We then define our problem and the model. After introducing the data set and the graph representation shaping our description, we proceed with presenting the results. This paper ends with a discussion on the results and an outlook on potential future research directions.

2 State of the art

Data-driven performance prediction approaches started early with works like [2], [3] to utilize ML methods such as long short-term memory (LSTM) to predict buildings thermal loads [4], [5], [6], [7], [8] or indoor climate conditions [9], [10], [11], [12], [13]. While all of these models achieved accurate results, they are still subjected to the problem of purely data-driven approaches: they often fail to make predictions complying with physical laws or lack the ability to generalize [1]. These shortcomings have motivated the development of physics-informed machine learning models like GNNs. GNNs are able to include physical knowledge into the learning process, since they are based on graphs. A basic graph consist of a set of nodes, connected by edges. Extending nodes with corresponding node features and correctly capturing causal, physical or spatial relations with edges, these structures offer a powerful way of encoding physical laws.

Various research has been conducted on employing GNNs for load prediction tasks on buildings with HVAC systems. [14] use a framework combining LSTM and GNN technology to provide heat load forecasting. Outdoor temperature, radiant flux, outdoor humidity content, and wind speed are used as features [15] to build a set of control relationship graphs based on spatial structure, control characteristics and equipment attributes and employ a GNN combined with an Informer model to predict an office buildings cooling load over an entire cooling season. [16] use a spatial temporal graph attention network to predict heating load for multiple thermal zones of an office building using outdoor features like solar radiation, temperature, hour of the day, wind speed and also indoor features like occupancy temperature differences of zones and between zones and outside.

Most of these approaches use a mixture of IoT data, outside climate influence and building usage data encoded in features like week day or hour of the day. Although these features are highly correlated to the target variable, we want to enable models to also deal with new situations. For example when there are significant changes in occupants behavior, time features quickly decrease in importance. To overcome this limitation and make use of a larger number of available sensors, we develop a model based on thermal zone sensors as well as HVAC internal data and outdoor climate.

3 Methodology

3.1 Problem definition

We define our approach as a graph learning problem on a building graph representation $\mathcal{G}$ that captures the thermodynamic relationships influencing the HVAC load.

Buildings thermodynamics vary over time, with some connections in the graph only being active in specific situations. Thus we use a temporal approach to capture the correct dynamics. In this approach, each timestep t_i has its own adjacency matrix A_{t_i}, which is derived from a timesteps graph $G_{t_i} = \{(\mathcal{V}_{stat}), (\mathcal{E}_{stat} + \mathcal{E}_{t_i})\}$. $\mathcal{V}_{stat}$ and $\mathcal{E}_{stat}$ are the static node and edge sets, shared across all timesteps and holding connections that do not change over time.$\mathcal{E}_{t_i}$ is a timesteps dynamically computed modified edge weight. $\mathcal{V}_{stat}$ denotes a set $\{v_1, v_2, ..., v_i\}$ of nodes representing a subset of all available sensors $\mathcal{V}_{stat} \subseteq \mathcal{S}$ in the data set. $\mathcal{E}_{stat}$ denotes a set of directed edges $\{e_1, e_2, ..., e_i\}$, each edge $e_{i,j} = (v_i, v_j)$ representing a connection directed from node v_i to node v_j.

By appending a feature vector $X_t \in \mathbb{R}^P$ to each node, with P the number of previous measurements, we receive the feature matrix $X \in \mathbb{R}^{(NxPxT)}$ with the number of nodes N and the number of timesteps T. Using the previously constructed adjacency matrices A, we define our TGNN as:

$$\tilde{y}_t = f(X, A_{t_i}) = \sigma(\hat{A}_{t_i}\, \mathrm{ReLU}(\hat{A_{t_i}} X W^{(0)}) W^{(1)}), \tag{1}$$

where $\hat{A_{t_i}} = \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}}$ is a preprocessing step, $\tilde{A_{t_i}} = A_{t_i} + I_N$ (with I_N the unity matrix) is a matrix that considers the features of the nodes for which the learning is conducted, $\tilde{D}$ is a degree matrix, where $\tilde{D} = \sum_j \tilde{A}_{ij}$, $W^{(0)}$ and $W^{(1)}$ represent the weight matrix in the first and second neighbourhood, and σ and ReLU represent sigmoid and ReLU activation functions.

3.2 Dataset and Scenario

To verify our approach, we utilize the data set presented in [17]. The data set offers time series data, ranging from January 2018 to December 2020. The data records contain energy use data outdoor and indoor environmental data and HVAC operational data. We will focus in particular on the data associated with the HVAC and the thermal conditions.

All data comes from a four floor building called Building 59, located in Berkeley, California. It contains 10.400 m^2 of air conditioned office space. The building is separated into two parts, the northern and the southern part with both parts containing 57 overall thermal zones. Since there is more data available for the southern part, we will use this part to build our models on.

The southern part of Building 59 is conditioned by two roof top air handling units (RTUs), serving the 20 thermal zones located in the building part. The loads of both RTUs are summarized to the HVAC load for the southern part.

In our scenario, $\mathcal{E}_{t_i}$ contains weighted connections between a thermal zone j's temperature and its cooling and heating setpoint. Both setpoints only influence a zone's climate if the corresponding system is active: the cooling setpoint if the HVAC fan is running, the heating setpoint if the heating valve is opened. Since there is also a quantitative difference in the influences depending on the valve opening/ fan speed, we map these values as the edge weights between the setpoints and the

zones temperature and scale them w_j to values $0 <= w_j <= 1$, with $w_j = 1$ indicating a zones fan is running at maximum speed or its valve is fully opened and $w_j = 0$ indicating a closed valve/not running fan. Computing this per timestep and thermal zone, we end up with the set of non-binary adjacency matrices A_{t_i}.

An interesting aspect of the data range from 2018 to 2020 is that it includes the Covid-19 pandemic. This poses an interesting test scenario for our prediction model: by training it on the data of 2019 and additionally testing it on the pandemic-influenced data, we can evaluate its ability to deal with out-of-training scenario. In this time, the usage pattern of Building 59 changed: the buildings occupancy and electrical energy usage decreased, but due to the low outside temperatures and a higher requirement of ventilation [17], the HVAC load does not decrease significantly. Figure 1 shows how the pattern of electrical usage change, as the occupancy drops and the consumption does not surge during the usual working hours.

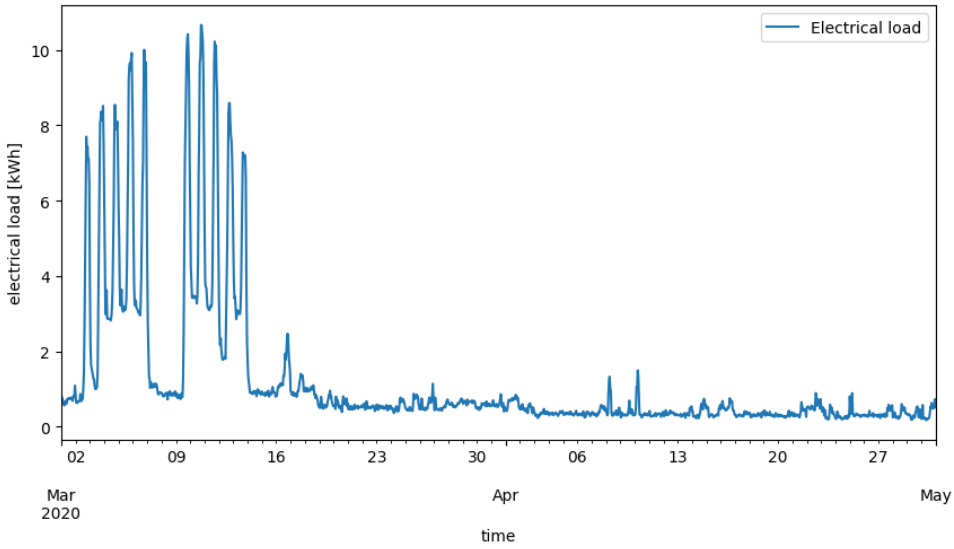


Figure 1: Hourly electrical energy usage at beginning of pandemic conditions

To additionally visualize that the pre-pandemic parts of the data set and the pandemic influenced parts differ and require a more causal model, we take a look at the Pearson correlation coefficients of features influencing HVAC load for both time series in Table 1. While outdoor environmental features like air temperature and relative humidity decrease in correlation, due to the higher ventilation requirements HVAC operational data gains influence on the load.

3.3 Graph representations and training

For the graph representation, we first encode the static part, constiting of nodes constantly contributing to HVAC load or influencing zone climate. In our scenario, this includes the sensors of both RTUs (003 and 004) as well as the fixed relations between the outside air temperature and the zones temperatures. Figure 2 shows a simplified visualization of the graph for RTU003.

The HVACs energy consumption is mainly governed by temperature difference between the supply air temperature and the return air and the fan speeds. We capture these relations in the graph. Using the various air temperature, setpoint and fan speed sensors inside the HVAC, we build the relations

Table 1: Pearson correlation coefficients for both parts of the data set

Feature	Pre-pandemic	Pandemic part
Outdoor air temperature	0.795	0.585
Outdoor relative humidty	-0.706	-0.341
RTU003 outdoor air damper position	0.137	0.354
RTU004 outdoor air damper position	-0.098	0.447
RTU003 supply fan speed	0.510	0.934
RTU004 supply fan speed	0.510	0.935

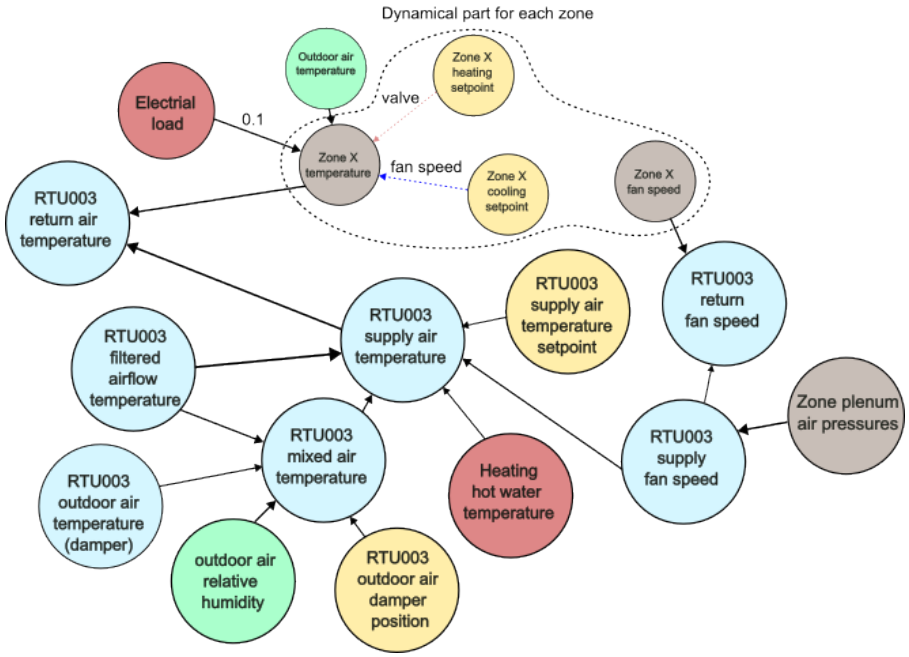


Figure 2: Graph representation for RTU003, including HVAC operational data (blue), setpoints (yellow), heat gains from electrical equipment and heating (red), thermal zone sensors (grey) and outdoor environmental data (green)

inside the HVAC. To include the influence of the thermal zones, we connect each zone to the return air temperature of its RTU. Because occupancy data is only available for short periods in the dataset, we use the electrial load as a substitution for internal heat gain. Since the internal gains do not influence a zone temperature in the same amount as other thermal relations, we weight this edge with a fixed value of 0.1.

Building up the static part of the graph for both RTUs and the thermal zones, we end up with 66 nodes connected by 107 edges. If all 20 thermal zones have opened valves and running fans, the number of edges can extend by 40 for a timestep. This approach gives us a solid representation of the thermophysics responsible for HVAC load.

The data is sampled in a 15 min interval. For training, we build a feature vector with $T = 8$ timesteps for each node, giving us a history of two hours. Our training data ranges from April 2019 to September 2019. Each feature vector is standardized using a standard scaler. We use a two layer TGCN for training, running with 50 epochs and optimized with the Adam optimizer, the learning rate is set to

0.01. A 80-20 % split is applied to separate training and test data (resulting in the pre-pandemic test data part).

4 Results

The results of our training are shown in Table 2. We use Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and R^2 as metrics to evaluate our models performance.

The model trained on the pre-pandemic data is able to maintain its performance, although decreasing in MAE and RMSE. This is also stated by the true versus predicted plot Figure 3. Although the correlations have shifted compared to the training, the model holds enough data points to not overfit on single features and anticipates the trend of the HVAC load data. It is especially remarkable, that the model accurately predicts load peaks, but struggles in periods with lower loads, underestimating the true values.

While R^2 score for the pre-pandemic data indicates a good regression of the problem, it significantly decreases on the out-of-distribution data.

Overall, the model demonstrates a strong ability to generalize from the high number of features it is given. Despite the decreasing performance on the pandemic data, it approximates the true values well and especially the trends and peaks. This indicates that the model has learned underlying patterns rather than overfitting to the training data or pure correlation, enabling it to adapt to previously unseen conditions.

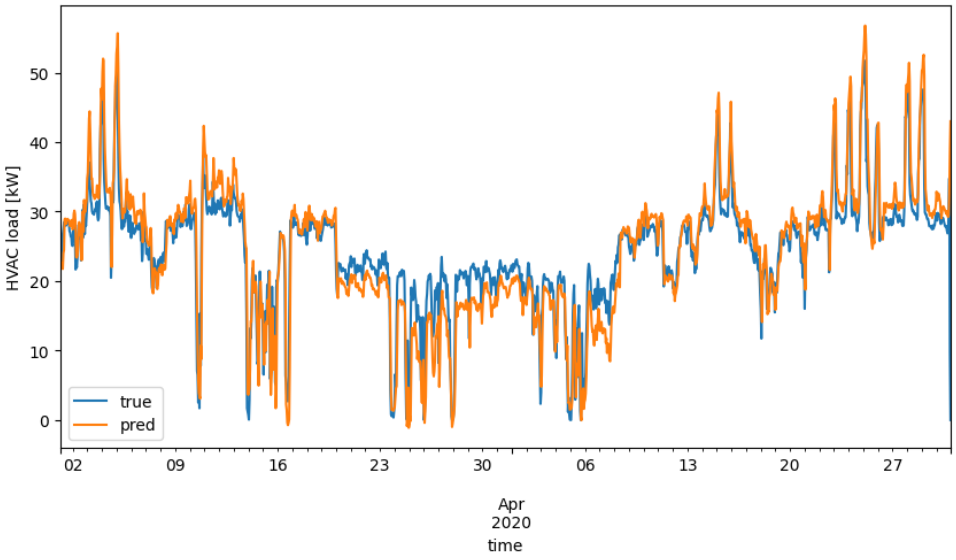


Figure 3: True vs predicted plot for pandemic data

Table 2: Model metrics for testdata and Covid 19 additional test data

Metric	Pre-pandemic test data	Pandemic test data
MAE	2.223	2.818
RMSE	3.049	3.817
R^2	0.941	0.776

5 Discussion and Outlook

The approach of incorporating a high number of IoT sensor into a temporal graph presented in this paper has shown to be resilient to changes in input data and maintain a high level of accuracy. Although the prediction accuracy falls short to other approaches presented in Section 2, the encoded physical knowledge lets the model stay on a similar level, even in unseen situations. Since machine learning approaches in practice often struggle to be accepted by decision-makers due to their black-box nature and their poor performance on unseen data [1], GNN/GCN approaches like the one presented in this paper are a step to more trustworthy predictions.

Despite the promising results, there is still a need for further research. As previously stated, interpretability of predictions also contributes to building trust in machine learning models. The approach presented here still remains a black-box, thus it is worth investigation the potential of explainability techniques for GNNs, like the ones presented in [18], [19].

Another area worth investigating is further automating the configuration of GNNs by automatically building the graph representation. Even tough the approach presented here requires less knowledge about the buildings characteristics than a physical simulation, a thorough domain knowledge on building physics is required. There have been effort to automate configuration using semantic math parsers [20] or LLMs [21], but there is not yet a reliable framework for the process.

However, the approach benefits from the high number of features, but these are not always available. The approach is yet to prove that it can be applied to other datasets than the one presented here.

References

[1] Z. Chen, F. Xiao, F. Guo, and J. Yan, “Interpretable machine learning for building energy management: A state-of-the-art review”, *Advances in Applied Energy*, vol. 9, 2023.

[2] A. E. Ben-Nakhi and M. A. Mahmoud, “Cooling load prediction for buildings using general regression neural networks”, *Energy Conversion and Management*, vol. 45, no. 13-14, 2004.

[3] Q. Li, Q. Meng, J. Cai, H. Yoshino, and A. Mochida, “Predicting hourly cooling load in the building: A comparison of support vector machine and different artificial neural networks”, *Energy Conversion and Management*, vol. 50, no. 1, 2009.

[4] J. Hu et al., “Thermal load prediction and operation optimization of office building with a zone-level artificial neural network and rule-based control”, *Applied energy*, vol. 300, 2021.

[5] S. Sholahudin and H. Han, “Simplified dynamic neural network model to predict heating load of a building using taguchi method”, *Energy*, vol. 115, 2016.

[6] Z. Wang, T. Hong, and M. A. Piette, “Building thermal load prediction through shallow machine learning and deep learning”, *Applied Energy*, vol. 263, 2020.

[7] A. Li, F. Xiao, C. Zhang, and C. Fan, “Attention-based interpretable neural network for building cooling load prediction”, *Applied Energy*, vol. 299, 2021.

[8] M. Zhou et al., “Issa-lstm: A new data-driven method of heat load forecasting for building air conditioning”, *Energy and Buildings*, 2024.

[9] F. Elmaz, R. Eyckerman, W. Casteels, S. Latré, and P. Hellinckx, “Cnn-lstm architecture for predictive indoor temperature modeling”, *Building and Environment*, vol. 206, 2021.

[10] N. Attoue, I. Shahrour, and R. Younes, “Smart building: Use of the artificial neural network approach for indoor temperature forecasting”, *Energies*, vol. 11, no. 2, 2018.

[11] B. Delcroix, J. L. Ny, M. Bernier, M. Azam, B. Qu, and J.-S. Venne, “Autoregressive neural networks with exogenous variables for indoor temperature prediction in buildings”, in *Building simulation*, Springer, vol. 14, 2021.

[12] E. Arulmozhi, J. K. Basak, T. Sihalath, J. Park, H. T. Kim, and B. E. Moon, “Machine learning-based microclimate model for indoor air temperature and relative humidity prediction in a swine building”, *Animals*, vol. 11, no. 1, 2021.

[13] N. Shakhovska, L. Mochurad, R. Caro, and S. Argyroudis, “Innovative machine learning approaches for indoor air temperature forecasting in smart infrastructure”, *Scientific Reports*, vol. 15, no. 1, 2025.

[14] X. Jia, H. Song, X. Nan, and X. Cai, “Prediction of short-term heat load of office buildings based on gnn-lstm modeling”, in *2024 6th International Conference on Frontier Technologies of Information and Computer (ICFTIC)*, IEEE, 2024.

[15] M. Zou, W. Huang, J. Jin, B. Hu, and Z. Liu, “Deep spatio-temporal feature fusion learning for multi-step building cooling load forecasting”, *Energy and Buildings*, vol. 322, 2024.

[16] Y. Jia, J. Wang, M. Reza Hosseini, W. Shou, P. Wu, and C. Mao, “Temporal graph attention network for building thermal load prediction”, *Energy and Buildings*, vol. 321, 2024.

[17] N. Luo et al., “A three-year dataset supporting research on building energy management and occupancy analytics”, *Scientific Data*, vol. 9, no. 1, 2022.

[18] Z. Ying, D. Bourgeois, J. You, M. Zitnik, and J. Leskovec, “Gnnexplainer: Generating explanations for graph neural networks”, *Advances in neural information processing systems*, vol. 32, 2019.

[19] H. Yuan, H. Yu, S. Gui, and S. Ji, “Explainability in graph neural networks: A taxonomic survey”, *IEEE transactions on pattern analysis and machine intelligence*, vol. 45, no. 5, 2022.

[20] A. Ba, K. Lynch, J. Ploennigs, B. Schaper, C. Lohse, and F. Lorenzi, “Automated configuration of heterogeneous graph neural networks with a semantic math parser for iot systems”, *IEEE Internet of Things Journal*, vol. 10, no. 2, 2022.

[21] D. Grauer and J. Ploennigs, “Does chatgpt know building physics? exploiting foundation models for building performance prediction with gnns”.

Prototypische Implementierung einer Konstruktions- und Regelsprache für Geometrie

Jan Hildebrand[1] and Marian Kurzawa[1]
[1]Fachgebiet Bauinformatik, Universität Kassel, Kassel, Deutschland
E-mail(s): j.hildebrand@uni-kassel.de, m.kurzawa@uni-kassel.de

Abstract: Digitale Bauwerksmodelle sind ein zentraler Bestandteil der Digitalisierung im Bauwesen. Sie werden unter anderem zur Überprüfung von normativen Anforderungen genutzt. Während explizite Informationen direkt aus Modellen extrahiert und kontrolliert werden können, sind geometriebasierte Prüfungen komplexer. Im Rahmen dieses Papers wird eine Konstruktions- und Regelsprache entwickelt, die neben der Erzeugung geometrischer Modelle (ähnlich wie in OpenSCAD) auch Funktionen zur Abfrage und Berechnung bauwerksspezifischer Parameter definiert. Diese orientiert sich an den Prinzipien formaler Sprachen. Für die Umsetzung der Konstruktions- und Regelsprache wird der Parsergenerator ANTLR v4 (ANother Tool for Language Recognition) herangezogen. Angestrebt wird ein Prototyp, mit dem ein Bauwerksmodell erzeugt werden kann und auf dessen Grundlage Abfragen zu Bauteileigenschaften sowie Berechnungen stattfinden können. Der Prototyp wird anhand von normativen Anforderungen, wie der Überprüfung von ausreichend Freiraum vor Türen, validiert.

Keywords: Regelprüfung, Geometriemodell, Algorithmen

DOI: 10.18154/RWTH-CONV-254883. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Einleitung

Digitale Bauwerksmodelle dienen der Speicherung, Kommunikation und Visualisierung bauwerksspezifischer Daten. Explizite Daten lassen sich unmittelbar aus dem Modell gewinnen, während geometrische Bauwerksprüfungen komplexere Verfahren benötigen. Zwar existieren Softwareprodukte mit Schnittstellen zur Implementierung entsprechender Prüfalgorithmen, jedoch erfordern diese Kenntnisse spezifischer Programmiersprachen. Die Relevanz dieser Problematik wird bereits in der Forschungsarbeit von Eastman et al. [1] betont. Die Autoren weisen darauf hin, dass die automatisierte Überprüfung von Bauwerksmodellen nach regulatorischen Anforderungen nach wie vor eine große Herausforderung darstellt. Als Lösung wird ein Prototyp einer Konstruktions- und Regelsprache entwickelt, die nicht nur geometrische Modelle, im Stile von OpenSCAD generiert, sondern auch bauwerksspezifische Parameterabfragen sowie -berechnungen ermöglicht. Während OpenSCAD einen deklarativen Ansatz verfolgt und beispielsweise keine Volumen in Variablen gespeichert werden können, soll dies im Prototyp möglich sein. Durch diese domänenspezifische Sprache soll die Hürde technischer Programmierkenntnisse reduziert und eine intuitive Möglichkeit geschaffen werden, um geometriebasierte Prüfungen nutzerfreundlich durchzuführen. In diesem Paper wird ein erster Prototyp

vorgestellt, der die Erzeugung von Bauwerksmodellen sowie deren regelbasierte Analyse ermöglicht und damit einen Beitrag zur praxisnahen Digitalisierung im Bauwesen leistet.

2 Methodik

Zur Entwicklung der Konstruktions- und Regelsprache werden zunächst theoretische Grundlagen formaler Sprachen, parametrischer Modellierung und des ANTLR-Frameworks erarbeitet. Die Grundsätze formaler Sprachen bilden die Basis für eine präzise Definition von Syntax und Semantik der Sprache. Die Eingabe des Nutzers werden in einem Geometriemodell erfasst und verarbeitet. Die parametrische Modellierung ermöglicht eine flexible Erstellung sowie Anpassung von Objekten im Geometriemodell. Das ANTLR-Framework bildet die Verbindung zwischen Sprache und Modell. Aus der definierten Grammatik erzeugt ANTLR einen Parser. Dieser gibt einen Syntaxbaum aus, welcher die Struktur der Eingabe abbildet. Um den Syntaxbaum zu traversieren, wird der in Java erzeugte Visitor verwendet. Der Visitor verarbeitet die einzelnen Knoten und übergibt die entsprechenden Objekte an das Geometriemodell zur weiteren Bearbeitung.
Im Rahmen einer Anforderungsanalyse werden konkrete Funktionen definiert, welche die Sprache unterstüzen: Erzeugen von Geometrien, Abfragen von Objektinformationen und Berechnungen auf Basis des Modells. Zur Syntaxgestaltung werden bestehende Konstruktions- und Regelsprachen, insbesondere OpenSCAD, als Referenz herangezogen, um eine intuitive Syntax abzuleiten. Der Prototyp orientiert sich an der deklarativen Syntax von OpenSCAD, der flexiblen Variablenverarbeitung ohne Typisierung und der modularen sowie kontrollstrukturbasierten Erweiterungsmöglichkeiten [2].
Die prototypische Implementierung der Konstruktions- und Regelsprache für Geometrie zieht die Ergebnisse ähnlicher Arbeiten heran, um Ansätze und Methoden in die eigene Entwicklung einzubeziehen: Lee et al. [3] betonen die Bedeutung einer klaren Trennung zwischen Regelformulierung, Modellvorbereitung, Regelausführung und Ergebnisdarstellung. Die Verwendung einer regelbasierten Sprache erhöht die Flexibilität von Prüfalgorithmen, da Regeln unabhängig von der konkreten Modellierungsumgebung formuliert werden können. Büttner et al. [4] zeigen, dass die Formalisierung von Regeln durch eine Grammatik und der Einsatz eines Parsergenerators wie ANTLR die Entwicklung einer erweiterbaren Modellierungssprache ermöglichen. Arias et al. [5] untersuchen die Anwendung von Constraint Logic Programming für Building Information Modeling und zeigen, dass logikbasierte Regeln und Constraints in die Modellierungsumgebung integriert werden können, um geometrische und semantische Zusammenhänge zu überprüfen. Pauwels et al. [6] liefern mit ihrem Ansatz zur semantischen Regelprüfung und Wissensrepräsentation im Bauwesen die theoretische und methodische Grundlage, warum formale, regelbasierte Sprachen für die Modellierung und Überprüfung von Bauwerksmodellen sinnvoll und notwendig sind. Zudem stellen Solihin und Eastman [7] eine Systematik zur Klassifikation von Regeln für automatisierte BIM-Regelüberprüfungen vor. Sie unterscheiden verschiedene Regelklassen, die unterschiedliche Komplexität und Anforderungen an die Modellierung und Abfrage von Bauwerksmodellen abdecken. Die erste Regelklasse umfasst einfache Prüfungen, bei denen einzelne explizite Eigenschaften auf Vorhandensein geprüft werden. Die zweite Regelklasse beinhaltet Bedingungen, die sich auf Beziehungen zwischen Objekten stützen. Die dritte Regelklasse bezieht sich auf Prüfungen, deren Berechnung Hilfskonstruktionen erfordert. Der Prototyp deckt Prüfungen bis zur dritten Regelklasse ab. Dies wird im Anwendungsbeispiel 4.2 geprüft.

3 Entwicklung des Prototypen

3.1 Grammatik

Für die Erstellung der Grammatik der entwickelten Konstruktions- und Regelsprache wird ein systematischer Ansatz gewählt, der auf den Prinzipien formaler Sprachen basiert. Zunächst werden die Anforderungen an die Sprache definiert, welche die Erzeugung geometrischer Objekte, das Abfragen von Eigenschaften sowie die Durchführung von Berechnungen umfassen. Aufbauend darauf erfolgt die Definition der lexikalischen Elemente (Tokens), welche die Grundbausteine der Sprache bilden. Anschließend werden syntaktische Regeln formuliert, welche die zulässige Struktur von Anweisungen und Befehlen festlegen. Dabei werden die Befehle in funktionale Kategorien unterteilt, „Create“ für die Objekterzeugung, „Inquiry“ für Abfragen und „Evaluate“ für Berechnungen. Befehle der Kategorie Edit dienen der aktiven Modellanpassung und sind durch ein „+“ gekennzeichnet. Sie umfassen das Erstellen (Create), Anpassen (Adjust) und Entfernen (Remove) von Objekten. Inquiry-Befehle hingegen sind mit „-“ markiert und führen keine Modelländerung durch, sondern dienen der Abfrage, Filterung, Anzeige von Eigenschaften, Berechnung (wie Fläche oder Volumen) und der Prüfung auf Überschneidungen zwischen Objekten. So wird eine klare Trennung zwischen modellverändernden und abfragenden Befehlen gewährleistet. Die Grammatik ist so aufgebaut, dass sie eine nutzerfreundliche Syntax bietet. Sie orientiert sich an bestehenden Sprachen wie OpenSCAD und ist formal maschinenlesbar. In Algorithm 1 sind jeweils ein Befehl für die Erstellung eines Objekts sowie für die Überprüfung dargestellt. Der Create-Befehl dient der Erstellung eines neuen Objekts einer bestimmten Klasse (wie door). Ein Objekt (Geometrie) wird über seine Position (pos) und die Ausrichtung (dir) definiert (siehe Abschnitt 3.2). Die Position entspricht dem Ausgangspunkt der Geometrie, der aus drei numerischen Werten besteht. Die Ausrichtung wird über eine Bezugsachse sowie einem Tripel aus positiven rationalen Zahlen angegeben. Optional können Parameter (props) des Objekts hinzugefügt werden. Hierfür muss ein Parametername (z.B. Material) und ein zugehöriger Wert angegeben werden. Mit dem *hasIntersection*-Befehl wird überprüft, ob zwei Objekte sich schneiden. Dafür müssen die IDs der beiden Objekte angegeben werden. Zudem kann mit dem Schlüsselwort „any“ eine Überschneidung mit sämtlichen Objekten des Modells abgefragt werden. Es wird ein Wahrheitswert sowie eine Liste von Elementen (bei Überschneidung) als Ergebnis zurückgegeben.

```
1    //Edit-Befehl
2    +new door (pos:(2,0,0), dir: x(1,0.375,2)), props("sustaining": true)
3
4    //Inquiry-Befehl
5    -hasIntersection wall1, wall2
```

Algorithm 1: Edit- & Inquiry-Befehl

Die Implementierung der Grammatik mit dem ANTLR4-Framework ermöglicht die Verarbeitung der Sprache im Java-basierten Prototyp. Die semantische Bedeutung der Befehle wird durch entsprechende Visitor-Methoden im Prototyp umgesetzt. Diese interpretieren die erzeugten Syntaxbäume und überführen sie in die Methoden des Geometriemodells. Abbildung 1 zeigt den Syntaxbaum des hasIntersection-Befehls aus Algorithm 1. Für das Modell existiert genau ein Programm (prog:1), das die oberste Hierarchieebene darstellt. Auf der darunterliegenden Ebene trennt der Parser sämtliche

Befehle (hier nur cmd:5). Befehle nach cmd:5 sind der Grammatik zufolge Anpassungen am Model. Ist der letzte Befehl der Eingabe erreicht, gibt EOF das Ende der Eingabe an. Im Beispiel wird das Token cmd:5 nach in der Grammatik formulierten Regeln weiterverarbeitet, bis die letzte Hierarchieebne erreicht ist. Hier sind die Token *hasIntersection*, *wall1* sowie *wall2*. Die Verarbeitung der einzelnen Token erfolgt über das Geometriemodell.

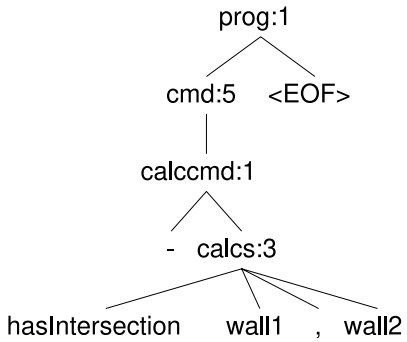


Abbildung 1: Beispiel Syntaxbaum

3.2 Geometriemodell

Das Geometriemodell ist objektorientiert konzipiert und ermöglicht die parametrische Modellierung von 3D-Objekten. Für den Prototyp sind einzig Solids als grundlegende geometrische Formen implementiert. Jedes Objekt im Modell besitzt definierte Attribute, die seine geometrischen Eigenschaften beschreiben, wie Abmessungen, Position und Orientierung im Raum. Darüber hinaus unterstützt das Modell die beschriebenen Operationen zur Anpassung und Abfrage dieser Eigenschaften. Dies ist essenziell für die spätere regelbasierte Prüfung und Berechnung.

```
1     public static void hasIntersection(String element1, String element2) {
2         // Wenn eines der Elemente "any" ist, pruefe Ueberschneidungen mit allen Elementen
3         if ("any".equals(element1)) {
4             checkIntersectionsWithAll(element2);
5         } else if ("any".equals(element2)) {
6             checkIntersectionsWithAll(element1);
7         } else {
8             // Normale Ueberschneidungsprüfung zwischen zwei Elementen
9             checkIntersection(element1, element2);
10        }
11    }
```

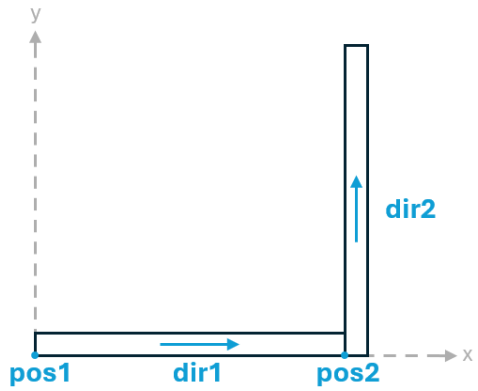
Algorithm 2: Visitor Methode: visitProg

Die Struktur des Geometriemodells ist so ausgelegt, dass eine beliebige Anzahl an Befehlen verarbeitet werden kann. Neu erzeugte Objekte werden in einer HashMap gespeichert. Für Anpassungen oder Berechnungen kann auf diese HashMap zugegriffen werden. Im Prototyp sind Methoden implementiert, die beispielsweise Schnittmengenberechnungen oder Abstandsmessungen zwischen Objekten erlauben. So können normative Anforderungen, wie die Überprüfung von ausreichend Freiraum vor Türen, geprüft werden. Die Verbindung von Grammatik und Geometriemodell stellt sicher, dass sowohl die Erzeugung als auch die Analyse der Bauwerksgeometrie konsistent und erweiterbar ist. Algorithm 2 zeigt beispielhaft die Methode zur Überprüfung auf Überschneidungen. Bezogen auf Abbildung

1 wird die Methode aufgerufen, wenn das Token *hasIntersection* erkannt wird. Die darauffolgenden Token *wall1* und *wall2* werden in den Strings *element1* und *element2* gespeichert. Die Prüfung auf Überschneidung erfolgt in ausgelagerten Methoden. Der Methodenkopf wurde von ANTLR generiert.

3.3 Einschränkungen

Die Entwicklung der Konstruktions- und Regelsprache und des dazugehörige Geometriemodell unterliegt mehreren Einschränkungen, die im Rahmen der prototypischen Implementierung berücksichtigt werden. Diese Einschränkungen gewährleisten, dass der Prototyp getestet und iterativ weiterentwickelt werden kann. Das Modell ermöglicht lediglich die Erstellung von achsparallelen, rechteckigen Geometrien, deren Ausrichtung sich an der X- oder Y-Achse orientieren. In Objekten hinterlegte Eigenschaften werden durchgehend als String gespeichert. Eine Integration bereits bestehender Bauwerksmodelle aus gängigen Formaten (z.B. IFC) ist bislang nicht umgesetzt, was die praktische Anwendbarkeit einschränkt. Zudem sind die automatisierten Prüfungen auf eine begrenzte Menge normativer Anforderungen fokussiert, wie die Validierung von ausreichend Freiraum im Türbereich. Die Nutzerfreundlichkeit des Prototyps ist durch das Fehlen einer grafischen Benutzeroberfläche noch eingeschränkt, da eine visuelle Rückmeldung fehlt, die Feedback zu den Eingaben liefert. Diese Einschränkungen bilden zugleich Ansatzpunkte für die zukünftige Weiterentwicklung und Erweiterung des Systems. In Abbildung 2 ist beispielsweise die Erstellung zweier Wände dargestellt. Für die prototypische Implementierung ist darauf zu achten, dass sich die Wände nicht überschneiden, sondern lediglich aneinander anschließen.




```
//new wall1
+new wall (pos: (0, 0, 0), dir: x(5.375, 0.375, 3))

//new wall2
+new wall (pos: (5, 0, 0), dir: y(5.375, 0.375, 3))
```

Abbildung 2: Erstellung von Objekten durch Koordinatendefinitionen

4 Test des Prototyps

4.1 Automatisierte Prüfung der Logik während der Modellerstellung

Bei der automatisierten Prüfung der Logik während der Modellerstellung soll sichergestellt werden, dass im Vorhinein festgelegte Regeln eingehalten werden, um die Korrektheit des Modells zu gewährleisten. Dazu gehört die Überprüfung der geometrischen Konsistenz, bei der beispielsweise sichergestellt wird, dass Objekte wie Wände sich nicht überschneiden. Hierdurch wird die Qualität des Modells verbessert und der manuelle Aufwand für die Fehlerbehebung reduziert. Die Überprüfung kann anhand unterschiedlicher Anwendungsbeispiele getestet werden. So kann zum Beispiel bei der

Erzeugung von Fenstern die Anforderung überprüft werden, ob das Fenster innerhalb einer Wand liegt und ob die Tiefen von Tür und Wand äquivalent sind. Sollte das Fenster diesen Anforderungen nicht entsprechen, gibt das Modell die Fehlermeldung „windowID is not fully contained within any wall.“ zurück. Analog kann bei der Erzeugung eines Raums überprüft werden, ob dieser von vier geschlossenen Wänden begrenzt wird. Dafür müssen Breite, Länge und Höhe des Raums mit der Begrenzung der Wände übereinstimmen. Die Methoden geben jeweils einen booleschen Wert zurück.

4.2 Validierung von ausreichend Freiraum vor einer Türöffnung

Im zweiten Beispiel wird geprüft, ob die Freiräume vor Türöffnungen den Vorgaben der DIN 18040-1 [8] entsprechen. Diese verlangt, dass die Bewegungsfläche vor Türen frei von Hindernissen wie Möbeln oder Einbauten sein muss und keine Stufen oder Schwellen aufweisen darf (siehe Abbildung 3).

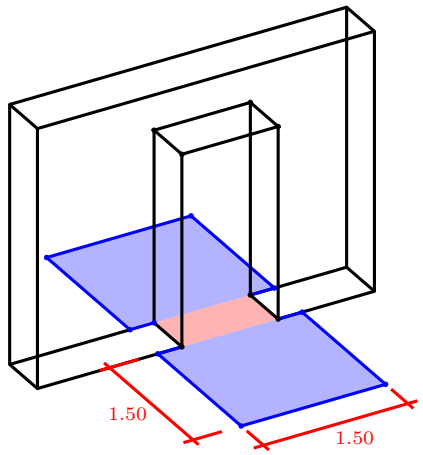


Abbildung 3: Freiraum vor Türöffnungen nach DIN 18040-1

Diese Regel entspricht der dritten Regelklassen, Berechnungen mit Hilfskonstruktionen, nach Solihin [7]. Der Prototyp besitzt noch keine Möglichkeit, Überschneidungen für bestimmte Objekte automatisch testen zu lassen. Die manuelle Erstellung eines Hilfskörpers (AuxiliaryBody) ist an der Stelle notwendig. Algorithm 3 ist der Eingabe zu entnehmen, dass neben den vier Wänden, die einen Raum erstellen, noch eine Tür, eine Säule sowie zwei Hilfsgeometrien erstellt werden.

Die Erstellung der Tür wird, wie in Abschnitt 4.1 dargestellt, auf korrekte Positionierung überprüft. Wird keine Fehlermeldung auf der Konsole zurückgegeben, befindet sich die Tür innerhalb der Wand. Die Säule ist so platziert, dass eine Überschneidung mit einem der beiden Freiräume (*auxiliarybody2*) auftritt. Die Hilfsgeometrien werden nach den Anforderungen der DIN 18040-1 erstellt, sodass sie direkt an die Tür/Wand anschließen. Mit dem Inquiry-Befehl „hasIntersection“ kann nun überprüft werden, ob die Hilfsgeometrien jeweils Überschneidungen mit anderen Objekten des Modells aufweisen. Der Algorithmus überprüft offene Intervalle, sodass die Ränder der Objekte rechnerisch keine Überschneidung aufweisen. In der Konsolenausgabe wird ersichtlich, dass für *auxiliarybody1* keine Überschneidung gefunden wird (siehe Abbildung 4). *Auxiliarybody2* weist hingegen eine Überschneidung mit der erstellten Säule auf. Somit entspricht der Freiraum vor der überprüften Tür nicht den Anforderungen der DIN 18040-1. Die Überschneidung der Hilfsgeometrie mit der Säule entspricht dem Ergebnis der Berechnung. Ebenfalls zeigt das Modell, dass der überprüfte Bereich lediglich bis

```
1  //Eingabe
2                                                                    //Variablenname:
3      +new wall (pos: (0, 0, 0), dir: x(5, 0.375, 3))               //wall1
4      +new wall (pos: (5, 0, 0), dir: y(5, 0.375, 3))               //wall2
5      +new wall (pos: (0.375, 5, 0), dir: x(5, 0.375, 3))           //wall3
6      +new wall (pos: (0, 0.375, 0), dir: y(5, 0.375, 3))           //wall4
7      +new door (pos: (2, 0, 0), dir: x(1, 0.375, 2))               //door1
8      +new column (pos: (1.0, 1.5, 0), dir: x(0.2, 0.2, 3))         //column1
9      +new auxiliarybody (pos: (0.5, -1.4, 0), dir: x(1.5, 1.4, 2)) //auxiliarybody1
10     +new auxiliarybody (pos: (0.5, 0.375, 0), dir: x(1.5, 1.4, 2)) //auxiliarybody2
11     -hasIntersection auxiliarybody1, any                          //INQUIRY (keine Var)
12     -hasIntersection auxiliarybody2, any                          //INQUIRY (keine Var)
13
14 //Konsolenausgabe
15
16     No intersections found for auxiliarybody1
17     Intersections found for auxiliarybody2: [column1]
```

Algorithm 3: Eingabe und Ergebnis des hasIntersection-Befehls

zur Türhöhe reicht. Dies kann relevant werden für Überprüfungen, bei denen die gesamte Raumhöhe benötigt wird.

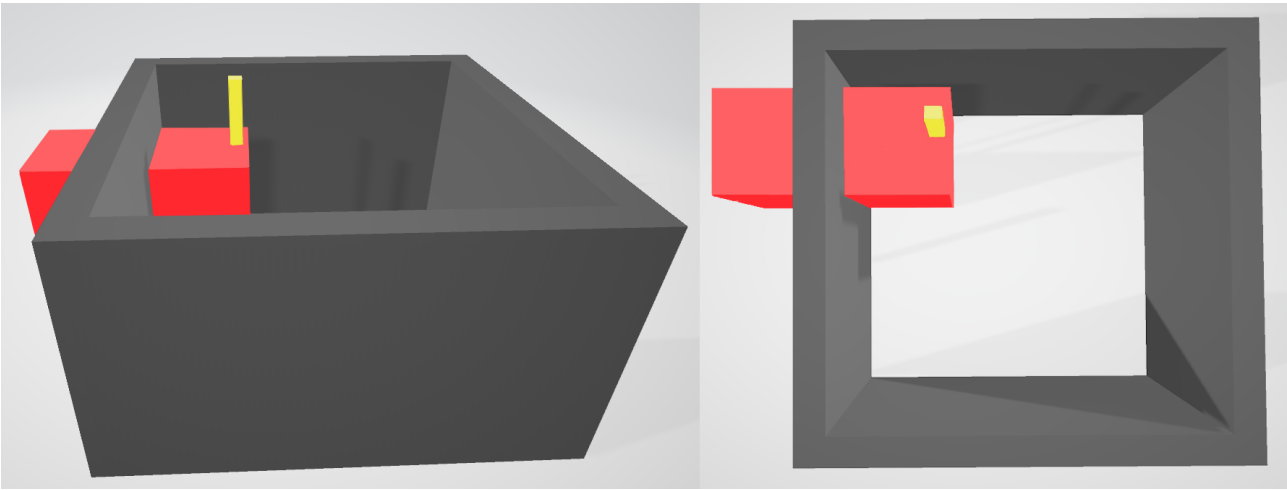


Abbildung 4: Darstellung des Türraums - Überschneidung mit Stütze vorhanden (Wände - grau, Hilfsgeometrien - rot, Säule - gelb)

5 Diskussion und Ausblick

Der entwickelte Prototyp der Konstruktions- und Regelsprache ermöglicht eine intuitive Modellierung und Überprüfung von Bauwerksgeometrien. Geometrische und alphanumerische Parameter können flexibel ver- und bearbeitet werden. Die Trennung von Befehlen zur Modelländerung und reinen Abfragen reduziert Fehlerquellen und unterstützt eine nutzerfreundliche Anwendung. Die Prüfung des Prototyps zeigt, dass sowohl die Abfrage und das Anpassen von Objekttatributen sowie die Überprüfung komplexerer geometrischer Anforderungen, wie die Prüfung des Freiraums vor Türen, möglich sind. Der Prototyp erleichtert den Zugang zu automatisierten Prüfungen im Bauwesen. Doch erfordert dieser für den praktischen Einsatz noch funktionale Verbesserungen.

Für die Weiterentwicklung des Prototyps ist primär die Erweiterung des Geometriemodells notwendig,

um komplexe und nicht achsparallele Geometrien abbilden zu können. Ebenso ist die Grammatik der Sprache so weiterzuentwickeln, dass funktionale Erweiterungen möglichst keine Änderung der Grammatik erfordern. Darüber hinaus sollte eine grafische Benutzeroberfläche implementiert werden, um die Bedienung ohne vertiefte Programmierkenntnisse zu ermöglichen. So können einzelne Modellierungsschritte vom Nutzer in Echtzeit nachvollzogen werden. Darüber hinaus bietet sich die Integration automatisierter Regelprüfungen an. So könnten normative Anforderungen bereits bei der Erstellung eines Bauwerkmodells überprüft werden.

Literatur

[1] C. Eastman, J.-m. Lee, Y.-s. Jeong und J.-k. Lee, »Automatic rule-based checking of building designs«, *Automation in Construction*, Jg. 18, Nr. 8, S. 1011–1033, 2009. DOI: https://doi.org/10.1016/j.autcon.2009.07.002

[2] OpenSCAD, *OpenSCAD User Manual/The OpenSCAD Language*.

[3] J. K. Lee u. a., »High-level implementable methods for automated building code compliance checking«, *Developments in the Built Environment*, Jg. 15, S. 100 174, 2023. DOI: https://doi.org/10.1016/j.dibe.2023.100174

[4] S. Büttner u. a., »Automatic scene parsing for generic object descriptions using shape primitives«, *Robotics and Autonomous Systems*, Jg. 76, S. 104, 2016. DOI: https://doi.org/10.1016/j.robot.2015.11.003

[5] J. Arias u. a., »Building Information Modeling Using Constraint Logic Programming«, *Theory and Practice of Logic Programming*, Jg. 22, Nr. 5, S. 723–738, 2022. DOI: 10.1017/S1471068422000138

[6] P. Pauwels u. a., »A semantic rule checking environment for building performance checking«, *Automation in Construction*, Jg. 20, Nr. 5, S. 506–518, 2011. DOI: https://doi.org/10.1016/j.autcon.2010.11.017

[7] W. Solihin und C. Eastman, »Classification of rules for automated BIM rule checking development«, *Automation in Construction*, Jg. 53, S. 73, 2015. DOI: https://doi.org/10.1016/j.autcon.2015.03.003

[8] Deutsches Institut für Normung e.V., *DIN 18040-1:2010-10, Barrierefreies Bauen – Planungsgrundlagen – Teil 1: Öffentlich zugängliche Gebäude*. Berlin, 2010.

Building a European-Level Energy Registry

Mahsa Faraji Shoyari[1], Jessy Matar[1] and Markus Duchon[1]
[1]fortiss GmbH, Munich, Germany
E-mail(s): farajishoyari@fortiss.org, matar@fortiss.org, duchon@fortiss.org

Abstract: The increasing integration of renewable energy across Europe demands a scalable and harmonized digital infrastructure to manage large volumes of energy-related data. This paper introduces the concept of a European Energy Registry (EER), aimed at establishing a framework that connects data providers, ensures semantic and structural integration of heterogeneous datasets, and exposes validated datasets to users. We examine the technical challenges in developing such a platform and outline a pipeline that leverages big data management, machine learning, and secure IT infrastructure. Key components include accurate data registration from diverse stakeholders, verification and correction mechanisms, and a robust, role-based user access system enhanced with multi-factor authentication and General Data Protection Regulation (GDPR) compliant security protocols. An API-driven data access model is proposed to enable seamless retrieval, filtering, and analysis of energy data. Automated validation and cleansing tools will support data quality and regulatory compliance. The underlying infrastructure will accommodate scalable databases, pipelines, and AI-driven analytics to provide predictive insights. Cloud-native deployment and APIs will ensure high availability, interoperability, and integration with external platforms. By addressing these requirements, the EER will support Europe's energy transition plan. This work presents a comprehensive technical blueprint for a secure, interoperable, and future-ready energy registry.

Keywords: Energy Registry, Data Management, Machine Learning, Interoperability

DOI: 10.18154/RWTH-CONV-254884. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Europe's transition to a decentralized and sustainable energy system has led to a surge in renewable energy integration. While supporting climate goals and market liberalization, this shift generates vast, fragmented, and heterogeneous datasets across national systems. Inconsistent standards and siloed data hinder cross-border collaboration, planning, and decision-making [1].

To address these challenges, a scalable, secure, and interoperable digital infrastructure is needed, one that supports seamless data exchange, integration, and validation across national and organizational boundaries. Recent advancements in big data management, machine learning (ML), and cloud-native technologies offer powerful tools to meet these requirements [2], [3].

A notable step in this direction is the NEED ecosystem [4], which lays the groundwork for a robust and flexible energy data space in the context of municipal and regional energy planning. NEED integrates

heterogeneous data sources via semantic models, employs synthetic data generation techniques, and exposes standardized APIs through a knowledge graph-based infrastructure. It emphasizes findability, accessibility, interoperability, and reusability (FAIR) data principles, data virtualization, and tool interoperability, demonstrating a scalable and modular approach to planning centric data integration.

Building upon the architectural and conceptual foundations of NEED, our work extends this vision to the European level by proposing a European Energy Registry, a cross-border digital infrastructure designed to standardize and streamline the access of energy data across member states. The EER aims to serve as a foundational component of the EU's energy data space, enabling trustworthy and interoperable data sharing among regulators, planners, researchers, and market actors.

Our proposed framework is aligned with the principles of Gaia-X and introduces additional energy-domain-specific capabilities. This work's key contributions include the introduction of a virtualization layer to enable on-demand, call-based data access, enhancing flexibility and reducing data duplication (1); the deployment of a workflow orchestration layer comprising automated data pipelines, cloud-native services, and monitoring mechanisms (2); the development of an OGC standard compliant spatial data access layer to support infrastructure planning and geospatial analytics (3); the integration of identity management, federated connectors, and secure access protocols to ensure data reliability and trust (4); the implementation of semantic interoperability tools such as ontologies, metadata catalogs, and knowledge graphs to link heterogeneous data sources meaningfully (5); and the use of synthetic data generation to address data availability and gaps where necessary (6).

The rest of the paper is structured as follows: Section 2 reviews related work; Section 3 presents the EER architecture; Section 4 outlines discussion and conclusion.

2 Related Work

The European energy sector struggles with fragmented data systems and complex regulations. Recent policy efforts and technological advances offer pathways to a more integrated and sustainable infrastructure. The EU Data Act (2024) enhances data transparency and usability, while the GDPR enforces strict rules on personal data use, critical for demand-side and retail energy applications. GDPR compliance is essential for building trust in AI-powered energy solutions [5]. To address these challenges, Europe is fostering sovereign, interoperable, and federated data spaces that enable secure, lawful, and transparent data sharing. Multiple initiatives are shaping this ecosystem through shared architectures and governance models.

2.1 Gaia-X and the International Data Spaces Association

Gaia-X is a European initiative for a federated, secure, and interoperable data infrastructure. In partnership with the International Data Spaces Association (IDSA), it supports cross-sector data ecosystems through domain-specific data spaces. IDSA contributes to the Dataspace Protocol, a modular architecture for secure, standards-based data exchange, emphasizing data sovereignty and semantic interoperability [6]. The IDS Information Model defines roles and policies, while the IDS Connector manages identity, contracts, and secure communication. Together, Gaia-X and IDSA enable scalable and trusted data ecosystems [7], [8]. Energy data-X, a Gaia-X lighthouse

project funded by BMWK, develops a federated energy data space supporting decentralized data exchange via standardized connectors and identity tools. Use cases include grid optimization and integration of flexible assets like EV chargers and heat pumps. OMEGA-X expands this vision at the European level, connecting energy domains-electricity, gas, mobility, smart cities through a secure, interoperable infrastructure. It promotes data sovereignty and cross-sector innovation with real-world pilots, supporting the energy transition and advancing a unified EU data space.

2.2 FAIR and Open Data

Germany's national research data infrastructure (Nationale Forschungsdateninfrastruktur - NFDI) fosters FAIR data practices like making data FAIR across scientific domains. The FAIR data spaces project expands this vision to industry, enabling secure, cross-sector data collaboration in line with Gaia-X principles. Use cases in health (NFDI4Health) and energy (NFDI4Energy [9]) demonstrate how distributed analytics and rich metadata can support innovation without compromising data sovereignty or privacy.

Complementing this, open data which is freely accessible and reusable by anyone promotes transparency, collaboration, and equitable access to knowledge, especially in resource constrained settings. While FAIR data doesn't necessarily mean "open," it ensures that data is shared under clear, well-defined conditions, balancing openness with the need to protect sensitive information.

Together, FAIR and open data approaches contribute to a federated, trustworthy, and inclusive data ecosystem [10].

While the above initiatives offer broad data-sharing frameworks, our blueprint addresses the European energy sector's complexity with advanced validation, AI analytics, and rich semantic metadata. This ensures a scalable, unified platform for efficient data management and the energy transition.

3 System Architecture of the EER

We present in this section, the core components that form the foundation of the proposed EER. We outline each component detailing the key building blocks of this framework starting from the southbound layer, followed by the middle and northbound layers (as seen in Figure 1 from bottom up).

3.1 Southbound Layer

The Southbound Layer serves as the primary interface for ingesting data from a wide range of heterogeneous distributed sources across the energy ecosystem. This layer must be designed to accommodate the diversity of formats, protocols, and update frequencies associated with different actors.

Aggregating different data sources forms the foundation for accurate analytics [11]. Among the primary energy data providers are Transmission System Operators (TSOs), who supply information on cross-border electricity flows, system frequency, and grid balancing operations. Distribution System Operators (DSOs) contribute detailed data at the local level, including voltage levels, transformer load status, and fault or outage events. Generation units, both renewable and conventional, feed the registry with production forecasts and output. At the consumption end, smart meters installed across residential, commercial, and industrial sites provide high-resolution usage data, enabling precise load profiling. In addition to these core contributors, energy market operators and exchanges provide

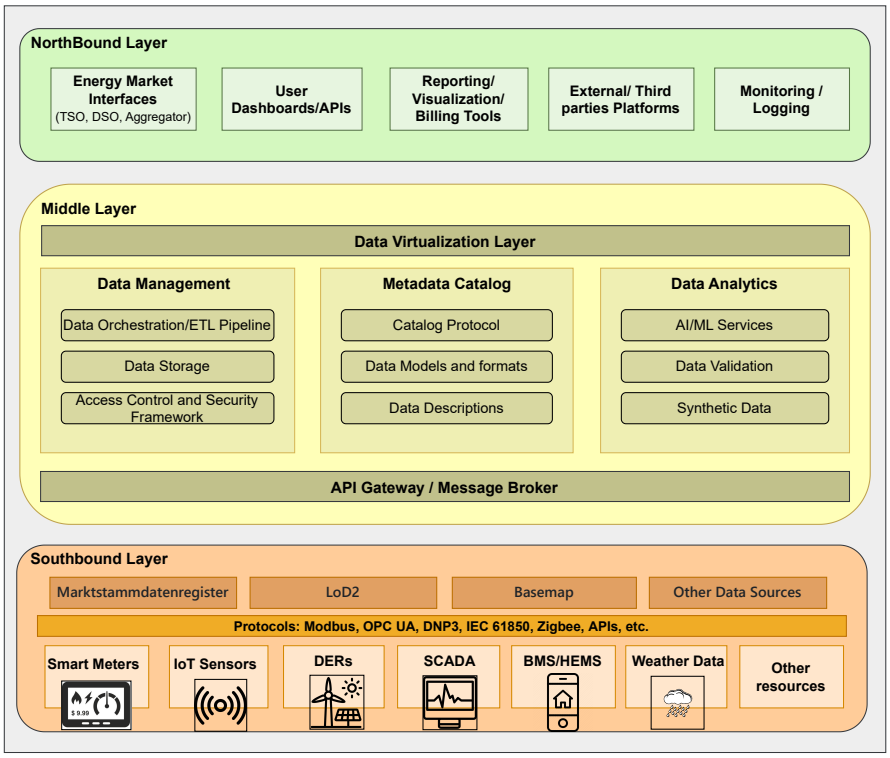


Figure 1: EER Components.

essential market data, such as pricing signals and transaction volumes from various trading platforms (e.g., day-ahead, intraday, and balancing markets). Meteorological data from weather services plays a key role in enabling accurate forecasting, particularly for variable renewable sources like wind and solar. Data from regulatory reporting systems and Virtual Power Plants (VPPs) also enrich the platform. This diverse and dynamic data landscape highlights the critical need for standardized data formats, secure communication protocols, and automated validation tools within the southbound interface of the EER, with adherence to OGC standards and the OGC API further enhancing interoperability across geospatial energy data sources.

3.2 Middle Layer

The Middle Layer acts as the processing core of the EER, responsible for tasks such as data connection and enrichment. It also manages metadata and ensures compliance with privacy and security standards.

3.2.1 Data Virtualization

Traditional data architectures, relying on centralized data lakes and warehouses, struggle to keep up with the complexity and distribution of modern data ecosystems. With data scattered across relational and NoSQL databases, graph stores, cloud filesystems, and SaaS platforms, centralizing all sources introduces redundancy, latency, and operational overhead. The EER addresses these challenges through a data virtualization approach. Rather than physically consolidating data, virtualization enables on-demand access to distributed sources through a unified interface [12]. This minimizes duplication, reduces latency, and ensures compliance. The platform architecture is structured around three key components: (1) Data Management, which handles orchestration, storage, and secure access, (2)

Metadata Catalog, which ensures semantic alignment and discoverability of data, (3) Data Analytics, which supports insights, validation, and synthetic data generation. To support these components, the platform emphasizes interoperability and scalability. Interoperability is essential for seamless data integration across systems. This project focuses on APIs, standards, semantic technologies, and Service Based Architecture (SBA) as key pillars. In parallel, containerization technologies like Docker and Kubernetes are fundamental to ensuring scalability and service accessibility. In the EER, Kubernetes orchestrates containerized components, automating deployment, scaling, and life cycle management.

1. **Data Management**

   Data Orchestration and ELT Pipelines In modern data infrastructures, orchestration and transformation tools are essential for building efficient and scalable data workflows. In the EER project, we adopt an ELT (Extract, Load, Transform) approach using Dagster and dbt (Data Build Tool). Dagster orchestrates the extraction process, connecting to diverse data sources and extracting only the required subsets, avoiding full data downloads and reducing resource overhead. dbt performs transformations directly within the target system, ensuring consistency and performance. This separation of concerns enhances pipeline transparency and simplifies maintenance while enabling distributed data integration.

   Data Storage While data virtualization minimizes the need to copy or replicate datasets, temporary or persistent storage is still necessary for analytical processing, caching, and performance optimization. Our architecture supports a hybrid storage model, including: (a) Relational and NoSQL databases for structured and semi-structured data, (b) Object storage solutions like MinIO for large binary or unstructured data, (c) Software-defined storage systems such as Ceph to ensure scalability and high availability in containerized environments. This flexible approach allows the platform to scale horizontally, handle diverse data types, and optimize cost-performance trade-offs across workloads.

   Access Control and Security Framework Given the presence of sensitive data especially in the southbound layer, robust access control and security mechanisms are critical [13]. Each data source enforces its own access policies, but the platform also implements a middleware security layer. We use Keycloak for centralized authentication and authorization, enabling role-based access control (RBAC) across services. This ensures that only authorized users and systems can interact with protected resources, in compliance with data protection regulations like GDPR.

2. **Metadata Catalog**

   Catalog protocol is central to the semantic, ontology driven framework upon which the metadata catalog of the EER would be based. This approach emphasizes structured vocabularies and relationships that formally define the meaning of data entities.

   Data models such as the CIM (Common Information Model), SAREF (Smart Appliances REFerence ontology), and the European Union's Smart Data Models enable consistent representation of concepts like energy generation, consumption, grid assets, and market roles across national

systems. Leveraging these models supports integration across heterogeneous energy datasets in Europe.

Data descriptions in the planned semantically enriched metadata catalog would use JSON-LD formatting and support SPARQL based querying and linked open data integration. This design allows datasets from different sources to be linked via shared concepts, facilitating automated data integration, validation, and compliance with regulatory frameworks.

3. **Data Analytics**
   Artificial intelligence and ML services can significantly enhance the utility of energy data. ML can go beyond traditional monitoring by learning complex patterns in consumption, generation, and grid conditions, enabling early detection of faults, more accurate renewable generation forecasts, and optimized balancing strategies. For example, ML-based predictive models can anticipate demand surges or wind power fluctuations hours ahead, allowing for proactive operational decisions that improve grid resilience.

   Data validation is necessary to ensure the accuracy and reliability of insights generated by ML models. Automated checks embedded in the data pipeline would assess completeness, consistency, and plausibility using a combination of rule-based logic, statistical profiling, and anomaly detection algorithms. In critical scenarios such as suspected sensor malfunctions a human-in-the-loop review can provide an additional safeguard, ensuring operational decisions are based on verified data.

   Synthetic data become important in scenarios where real-world data is scarce, incomplete, or sensitive. Techniques such as Generative Adversarial Networks (GANs) or agent-based simulations [4] can generate such data. Synthetic datasets also support cross-border experimentation by providing standardized test cases that reflect the statistical properties of actual energy datasets [14].

3.2.2 API Gateway/Message Broker

To support secure and efficient communication in a distributed, service-based architecture, the platform integrates both an API Gateway and a Message Broker. The API Gateway acts as a single entry point for external and internal API requests, handling authentication, routing, rate limiting, and protocol translation, ensuring secure and consistent access to microservices. Meanwhile, the Message Broker (e.g., Kafka or RabbitMQ) facilitates asynchronous, decoupled communication between components, enabling high-throughput data exchange, buffering, and event-driven processing.

3.3 Northbound Layer

The Northbound Layer serves as the interface between the energy data platform and end-users. It provides user-facing tools like dashboards, APIs, and visualization services, enabling seamless access to processed and validated data. Additionally, it facilitates integration with third-party platforms and supports monitoring and logging functionalities. This layer ensures that insights derived from underlying data systems are securely and efficiently delivered to relevant stakeholders for decision-making and operational use.

4 Discussion and Conclusion

The EER is an ambitious project that comes with a wide range of challenges. Designing and implementing a registry that can handle Europe's diverse energy data landscape is highly complex, from data integration to security and governance, to long-term scalability. Each energy actor generates data in its own format, often using different communication standards. Harmonizing such inputs requires sophisticated ETL pipelines and hybrid storage architectures that can evolve over time. Much of the data in question relates to critical infrastructure or sensitive operational behavior. The EER must enforce strict data protection through end-to-end encryption to ensure transparency and accountability. In addition, energy data is prone to inconsistencies, such as missing values, and errors stemming from device faults or communication failures. Hence, it is critical to implement robust ML integration for validation and data cleansing. Finally, the volume, velocity, and variety of energy data are expected to grow rapidly as more distributed energy resources come online. The EER must therefore adopt cloud-native architectures and scalable data processing pipelines to handle these future demands while ensuring low latency and high availability across a European infrastructure.

The proposed European Energy Registry represents an important step towards creating a more integrated and transparent energy data ecosystem across the EU based on a detailed reliable metadata catalog. As we have seen, the path to implementing the EER involves a high degree of complexity. Ultimately, the EER should be viewed as a long-term initiative that must remain flexible and open to future innovations in data management, artificial intelligence, and grid technologies. As the European energy landscape continues to evolve, so the infrastructure that supports it. The framework outlined in this paper provides a foundation for such an effort, but much work remains to turn this vision into a practical and widely adopted system.

Acknowledgements

We gratefully acknowledge financial support through the project executing agency Jülich (PTJ) with funds provided by the Federal Ministry for Economic Affairs and Climate Action (BMWK) due to an enactment of the German Bundestag under Grant No. 03EN3077%.

References

[1] R. Verzijlbergh, L. Vries, G. Dijkema, and P. Herder, "Institutional challenges caused by the integration of renewable energy sources in the european electricity sector", *Renewable and Sustainable Energy Reviews*, vol. 75, Nov. 2016. DOI: 10.1016/j.rser.2016.11.039

[2] G. Rostirolla et al., "A survey of challenges and solutions for the integration of renewable energy in datacenters", *Renewable and Sustainable Energy Reviews*, vol. 155, no. C, 2022. DOI: 10.1016/j.rser.2021.111787 [Online]. Available: https://ideas.repec.org/a/eee/rensus/v155y2022ics136403212101056x.html

[3] S. Onwusinkwue et al., "Artificial intelligence (ai) in renewable energy: A review of predictive maintenance and energy optimization", *World Journal of Advanced Research and Reviews*, vol. 21, pp. 2487–2799, Jan. 2024. DOI: 10.30574/wjarr.2024.21.1.0347

[4] M. Duchon et al., “A platform ecosystem providing new data for the energy transition”, *SIGENERGY Energy Inform. Rev.*, vol. 4, no. 4, pp. 226–237, Feb. 2025. DOI: 10.1145/3717413.3717435 [Online]. Available: https://doi.org/10.1145/3717413.3717435

[5] B. Jørgensen and Z. Ma, “Regulating ai in the energy sector: A scoping review of eu laws, challenges, and global perspectives”, *Energies*, vol. 18, p. 2359, May 2025. DOI: 10.3390/en18092359

[6] A. Braud, G. Fromentoux, B. Radier, and O. Le Grand, “The road to european digital sovereignty with gaia-x and idsa”, *IEEE network*, vol. 35, no. 2, pp. 4–5, 2021.

[7] S. Bader et al., “The international data spaces information model–an ontology for sovereign exchange of digital content”, in *International Semantic Web Conference*, Springer, 2020, pp. 176–192.

[8] J. Pampus, B.-F. Jahnke, and R. Quensel, “Evolving data space technologies: Lessons learned from an ids connector reference implementation”, in *International Symposium on Leveraging Applications of Formal Methods*, Springer, 2022, pp. 366–381.

[9] M. Faraji Shoyari and M. Duchon, “Scalable and interoperable data infrastructure in energy domain”, in *2. NFDI4Energy Conference 2025*, 2025. [Online]. Available: https://doi.org/10.5281/zenodo.14961100

[10] J. Theissen-Lipp et al., “Semantics in dataspaces: Origin and future directions”, in *Companion Proceedings of the ACM Web Conference 2023*, 2023, pp. 1504–1507.

[11] D. M. Dave and B. Mittapally, “Data integration and interoperability in iot: Challenges, strategies and future direction”, *INTERNATIONAL JOURNAL OF COMPUTER ENGINEERING & TECHNOLOGY*, vol. 15, pp. 45–60, Jan. 2024.

[12] O. Oyeniran, O. Modupe, A. Otitoola, O. Abiona, A. Adewusi, and O. Oladapo, “A comprehensive review of leveraging cloud-native technologies for scalability and resilience in software development”, *International Journal of Science and Research Archive*, vol. 11, pp. 330–337, Mar. 2024. DOI: 10.30574/ijsra.2024.11.2.0432

[13] C. Vigurs, C. Maidment, M. Fell, and D. Shipworth, “Customer privacy concerns as a barrier to sharing data about energy use in smart local energy systems: A rapid realist review”, *Energies*, vol. 14, p. 1285, Feb. 2021. DOI: 10.3390/en14051285

[14] J. Matar, M. Koubaa, and M. Duchon, “Exploring the potential of freely available satellite data for energy applications”, in *2. NFDI4Energy Conference 2025*, 2025. [Online]. Available: https://doi.org/10.5281/zenodo.14843602

Transforming IFC Models into a Structured 3D Format for Multimodal Localization from 2D Images in a Web-Based Tool

Robin Ov[1], Ayman Soultana[1], Angelina Aziz[1] and Roberto Perez Martinez[2]
[1]Ruhr University Bochum (RUB), Bochum, Germany
[2]Fraunhofer Heinrich Hertz Institute (HHI), Berlin, Germany
E-mail(s): robin.ov@ruhr-uni-bochum.de, ayman.soultana@ruhr-uni-bochum.de, angelina.aziz@ruhr-uni-bochum.de, roberto.perez.martinez@hhi.fraunhofer.de

Abstract: Accurate localization within built environments based on 2D images remains a fundamental challenge in construction informatics, robotics, and augmented reality applications. Existing approaches, such as SPVLoc (Semantic Panoramic Viewport Matching for 6D Camera Localization), have demonstrated reliable image-based localization by leveraging structured 3D representations. However, many localization pipelines primarily rely on point clouds and structure-from-motion, which lack the semantic richness of Building Information Modeling (BIM) data. This study investigates the transformation of Industry Foundation Classes (IFC) models into a Structured3D format, a representation that combines 3D geometric primitives with semantic relationships, to improve localization accuracy using a multimodal approach and its integration into a web-based tool. Structured3D provides localization algorithms with contextual information beyond raw geometrics, enabling more robust pose estimations in complex indoor environments. The conversion from IFC models into a Structured3D format is implemented within a web-based tool that incorporates an IFC viewer, allowing users to interactively visualize and process the Structured3D data. The localization workflow integrates image-based and geometric data, enabling precise pose estimation. Camera positions obtained from the localization process are stored within the IFC model as documentation points, providing a structured record of captured viewpoints for further analysis. The approach is evaluated in real-world building environments, considering varying lighting conditions, occlusions, and perspective changes. The results indicate that IFC-derived Structured3D data, combined with web-based integration, enhance localization accuracy and usability. Future research will explore ML services to automatically detect building components and estimate their position within this pipeline.

Keywords: IFC, Structured3D, localization,SPVLoc, BIM, multimodal approach, 2D-to-3D matching, web-based tool

DOI: 10.18154/RWTH-CONV-254885. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Accurate indoor camera localization within architectural environments is a fundamental challenge in fields such as navigation, augmented reality, building management and robotics. Traditional 2D-to-3D matching techniques, from 2D camera images to locations in a 3D model, often rely on point clouds or on CAD-based models, which are geometrically detailed, but lack semantic notations for a comprehensive understanding of the scene. This lack of semantic information hinders effective localization, especially in complex indoor settings.

Research has shown the potential of integrating semantic information in localization pipelines. For example, Structured3D provides a dataset of synthetic annotated indoor scenes in which both geometric primitives and their semantic interpretation are described within a single framework [1]. This allows for detailed modeling of complex indoor environments, capturing both the arrangement of objects and their defined grouping. In turn, Structured3D has shown improved performance on Machine Learning tasks requiring holistic understanding, such as room layout estimation.

However, while Structured3D provides a powerful framework for scene representation, its reliance on synthetic scenes limits its applicability to real-world architectures. In practice, semantically rich models of buildings are ideally stored in Industry Foundation Classes (IFC), a standard format used in Builidng Information Modeling (BIM). However, the structure of IFC is complex and difficult to directly apply to ML tasks [2]. As a result, we propose a converter that transforms IFC models into the Structured3D format. By extracting IFC entities and their geometric primitives and reencoding them in the Structured3D schema, our tool enables a structured semantic representation of real-world architectural environments, enabling compatibility with ML tasks. We will also apply our converted S3D json files to SPVLoc, which is a camera localization method, based on the S3D framework[3].

2 Background

Indoor camera localization is a key task in computer vision that involves a range of 2D to 3D localization techniques, leveraging point clouds, CAD models, and semantically annotated datasets. The following section proves a detailed overview of the background on indoor camera localization.

Point cloud camera localizations like PICCOLO make use of both colored point clouds and a query panorama image of a scene for localization [4].This approach enables localization by matching the query image to the point cloud using a gradient descent algorithm, without requiring neural network training or the collection of ground-truth image poses. Other methods using point clouds along with image queries for camera localization include line-based 2D-3D pose estimation approaches that leverage geometric features in untextured point clouds, by extracting line segments [5]. In contrast, i3dLoc adresses inconsistencies between query images and point clouds, due to conditions such as lighting or weather [6]. Their method leverages a Generative Adversarial Network to extract condition-invariant geometric features to match these query images with point clouds. But all of these methods share the limitation of the requirement of pre-scanning all rooms on which localization should be conducted, which can be costly, especially in large buildings.

Consequently, the advantages of utilizing a common format for capturing environment information become apparent. One widely adopted standard in architecture is the Industry Foundation Classes

(IFC) format, which is a general-purpose data format designed to enable interoperability across different software [7]. But the feature of utilizing IFC for different software comes at the cost of a complex structure, which are often underutilized in specific use cases. As a result, extracting relevant information can be inefficient and challenging, requiring conversion methods to tailor the data to the targeted application.

While IFC offers a comprehensive standard for building data, its complexity can limit its direct applicability in certain tasks. To better address specific needs, several specialized formats have been developed, such as CityGML for urban-scale 3D city representations, GeoJSON for geospatial data exchange or point clouds for detailed 3D scanning and mapping [8], [9], [10]. For representing indoor environments in camera localization, we use the Structured3D json format, which addresses the lack of comprehensive information for training ML models [1]. The format of Structured3D makes use of geometric primitives such as **planes**, **lines** and **junction points**. The relationships of these primitives are also included in the format by matrices, defining their connectivity to define larger geometric objects along with their semantic interpretation. Environments in Structured3D format can currently be generated from floorplan images [11]. But the conversion can include errors due to low image qualities. Using IFC files comes with the benefit of being less vulnerable to errors, as IFC files contain detailed classes specifically defined for structural elements, such as walls, doors and windows.

In previous studies, the Structured3D format has been used by SPVLoc to perform camera localization by combining the format with images [12]. Structured3D allows SPVLoc to avoid the need of detailed textured models or dense point clouds, cutting on costs. The pipeline involves utilizing a perspective RGB images to be localized to rendered semantic panoramas from Structured3D. Important to note is, that SPVLoc does not require fine-tuning for unseen Structured3D environments, enabling a quick and scalabe deployment which was successfully demonstrated in an ML-based visual fire safety inspection pipeline [11]. Similar 3D-based localization, such as Inloc rely on dense, image-derived 3D models without semantic abstractation and require detailed scans of the environment [13]. Other methods focus on specific application domains, such as Lost Shopping!, which performs camera localization from images within large shopping malls, requiring query images and floor plans as input [14]. Or Rent3D, which focuses on estimating room layouts from single images using floor plan priors, but do not perform global localization [15]. In contrast, SPVLoc offers a general-purpose solution that can operate across diverse indoor environments without scene-specific adaptation [3].

3 Methodology

The following sections outline the data format of Structured3D and the conversion process of ifc to Structured3D, which enables the localization process by SPVLoc.

3.1 Technical Background

IFC is a standardized object-based schema defined by entity hierarchies. The entities inherit properties of their respective parent entities, such as IfcWall inherits from IfcBuildingElement. The geometry are defined by boundary representations, such as faces, edges, and vertices. The format of Structured3D is a JSON dictionary, which makes use of geometric primitives such as **junction**, **lines** and **planes** points to define their relationships as seen in Figure 1.

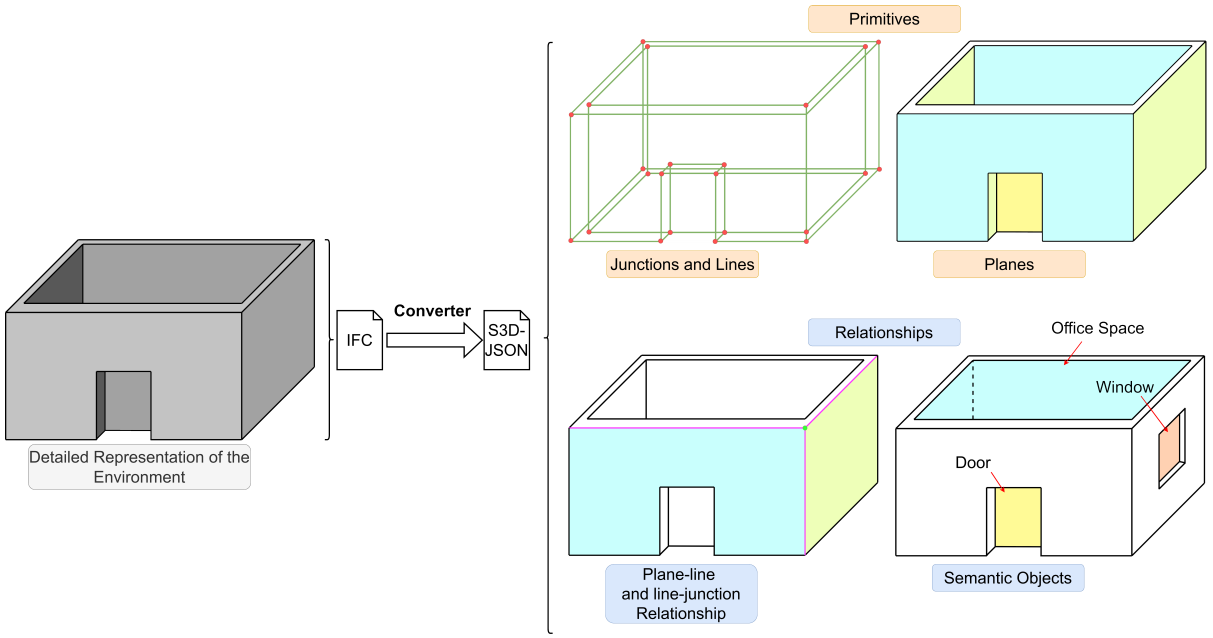


Figure 1: Overview of the IFC to S3D converter

As shown in Figure 1, the **junction points** are defined as individual 3D coordinates. While the **lines** are defined by the coordinate of a junction point and the direction to the other connected junction. **Planes** are defined by their normal and offset. The **lineJunctionMatrix** is of dimension LxJ, where L and J represent the lines and junctions, respectively. Each entry is of binary value, such that lineJunctionMatrix$_{lj} \in \{0, 1\}$. The value defines if the corresponding line and junction is connected. The **planeLineMatrix** follows the same structure as the lineJunctionmatrix, but it instead defines if the plane is bounded by the corresponding line. The **semantics** are defined by a tuple of planes, with the according name of the object. The initially proposed Structured3D format also contains cuboids and manhattan, however they were omitted in this work, as they are not utilized by the SPVloc pipeline [1].

3.2 Conversion of IFC

As a first step, the IFC model's units are converted to millimetres to standardize all geometric measurements. The conversion process is divided in four main sections, extraction of lines and junctions, extraction of planes, extraction of semantics and openings, and finally the generation of the line-junction and plane-line matrix, as they can be seen in the previous Figure 1.

3.2.1 Extraction of Lines and Junctions

The first step involves parsing the IFC geometry to extract junction points and lines from the geometric mesh representations of building elements. The IFC building elements include IfcWall, IfcDoor, IfcWindow, IfcColumn, and IfcSpace. Each object is converted into a mesh using the ifcopenshell library, after which the vertices and edges are extracted. The 3D coordinates of the vertices are registered as junctions. The edges between the vertices are registered as lines. Objects that fill openings of walls, such as IfcDoor and IfcWindow, are saved seperately, as they are later used in the semantics section to annotate opening objects.

3.2.2 Extraction of Planes

The second step focuses on extracting planes from the building geometry. For each IfcWall, polygonal faces are identified and transformed into global coordinates. Planes are then defined using the face normal and offset, computed from three non-collinear points. Only vertical faces of walls are defined as planes, due to their horizontal faces not being visible from inside the rooms. The IfcRelVoidsElement relationship is used to identify openings, such as doors and windows, and associate them with the appropriate wall geometry. This entity in the IFC schema relates an opening element (such as an IfcOpeningElement) to its host wall. During parsing, each IfcRelVoidsElement is checked to ensure the voided element is an IfcWall, and a mapping from opening ID to wall ID is created. Once the corresponding wall is found, its associated vertical planes are retrieved. Next, the geometric mesh of the IfcOpeningElement is extracted, and its vertices are converted into millimetre-scale coordinates. From the bounding box of the opening, the shortest axis is determined, which is assumed to represent the depth or thickness of the opening (i.e., the direction across the wall). This axis is used to split the opening geometry into two parallel surfaces (front and back), which represent the two visible faces of a door or window opening in the wall. For both faces, a convex hull is computed over the corresponding face vertices to determine the lines of the polygon. Line directions are determined by traversing the loop, generating 3D lines defined by their starting junctions and directions. These lines are then evaluated against candidate wall planes to identify which wall plane each corresponding opening face lies on. This step is crucial, as Structured3D expects wall planes to include the door lines, which are assigned later in the planelinematrix. Finally, these front and back faces of the opening are assigned a new plane with the same parameters as their corresponding wall plane. Each vertex is registered as a junction, and the edge direction vectors are stored as lines. As a post-processing step, the normals of each plane are adjusted to face outward from the corresponding IfcSpace, which is the room they belong to. In addition to wall and opening surfaces, the floor and ceiling are also added as planes, based on the horizontal top and bottom plane of the IfcSpaces.

3.2.3 Semantics

In the next step, the semantic structure of the IFC is extracted. This includes identifying all rooms and assigning their corresponding wall,floor and ceiling planes to each of them. Doors and windows are also added as seperate semantic elements. Each door or window is assigned the same ID as the room it connects to, meaning they are treated as part of that room's description.

3.2.4 Plane-Line and Line-Junction Matrix

The final components needed for the Structured3D JSON are the Plane-Line and Line-Junction matrices. The Plane-Line matrix specifies which lines lie on which planes. For each object, all lines are evaluated against each of its planes, and lines on the plane are marked 1; otherwise 0. For doors and windows, lines must also be registered on the corresponding wall planes if they align. The Line-Junction matrix links each line to its defining junctions. For every line, the two nearest junctions corresponding to its start and end are identified within the object it belongs to and marked in the matrix.

With all the IFC information extracted, the Structured3D JSON is created. It includes sections for all junctions, planes, semantics, and the matrices. This JSON can now be used by SPVLoc to perform localization. Next, the results and a look of the implementation in the web-based tool will be shown.

4 Results

This section presents the implementation and evaluation of the IFC to Structured3D conversion pipeline integrated with SPVLoc. The results demonstrate the conversion interface and an example of localization performance using real building data.

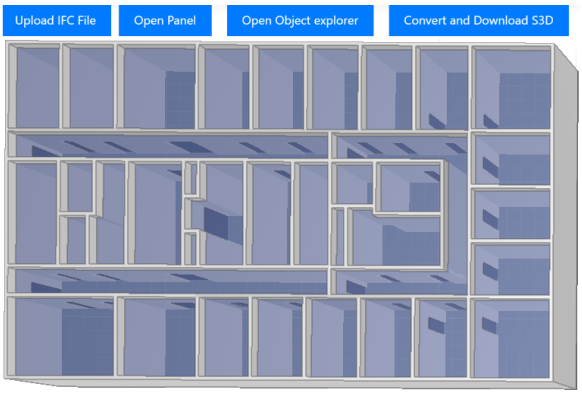


Figure 2: Web based tool for uploading an IFC file and converting it to Structured3D

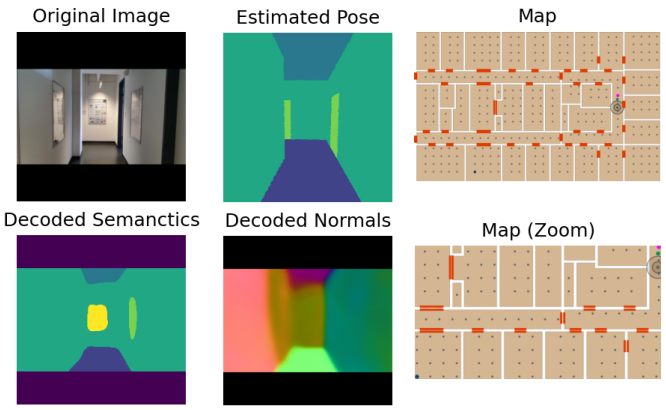


Figure 3: Exemplary result from SPVLoc

Figure 2 shows the website interface developed for the conversion pipeline. It allows users to upload an IFC file, explore the object hierarchy and properties, and convert the file into a Structured3D JSON format. For this work, an IFC model of the IC6 floor at Ruhr University Bochum was used as input. The resulting Structured3D JSON file was then used in SPVLoc, along with corresponding images, to perform localization within the environment. Figure 3 shows an exemplary result of this pipeline, detailing each stage of the SPVLoc localization process. The top-left image shows the original RGB input image captured in the hallway. This image is processed through the SPVLoc pipeline to infer semantic segmentation labels and surface normals, depicted in the decoded semantics and decoded normals views, respectively. The estimated pose is shown in the top-middle image, and uses these intermediate to predict the camera orientation and location relative to the building's Structured3D layout. The rightmost images illustrate the localization from a top-down view of the environment. The green dot indicates the predicted camera position, the pink dot shows the estimated orientation of the camera, and the grey dot with the concentric cirles marks the ground truth position derived from precise annotations. The second best matching location is visualized by a dark blue dot. A zoomed-in view of the map is also included to focus on these points of interest. These results highlight how the converted Structured3D data serves as spatial and semantic reference, enabling accurate and interpretable localization within the built environment.

5 Discussion

While the conversion of IFC model to the Structured3D format enables accurate localization, several practical challenges were encountered. The planeLineMatrix and lineJunctionMatrix, encode binary

entries which can lead to large json files due to their quadratic growth as the number of planes, lines, and junctions increases. To mitigate file size, index-based tuples have been implemented to represent connections more compactly. Another limitation arises from the use of mirrored or improperly aligned IFC models, which can result in inaccurate localizations. Ambiguity in semantic mapping can also be a challenge, especially when it's unclear in the IFC model whether an element is an IfcSpace (i.e., a room) or not. Despite these issues, the IFC-to-S3D conversion proved effective, as Structured3D representations were consistently generated, and localization based on them were accurate. The results demonstrate that 3D data extracted from IFC can successfully be used for camera localization or other applications which require spatial reasoning.

6 Conclusion

This work presents a complete pipeline for converting IFC models into the Structured3D format by leveraging the rich information available in IFC files. This enables effective localization, as demonstrated using SPVLoc. For future work, it would be valuable to address challenges such as mirrored IFC models, where identical looking areas in a building can lead to ambiguous localization. In the future, techniques that progressively refine uncertain semantic assignments or adapt to model inconsistencies could help improve robustness in such cases [16]. Extending the pipeline to include additional object types from the IFC model, such as equipment or furniture, could further help resolve ambiguities in localization.

References

[1] J. Zheng, J. Zhang, J. Li, R. Tang, S. Gao, and Z. Zhou, "Structured3d: A large photo-realistic dataset for structured 3d modeling", in *Computer Vision – ECCV 2020*. Springer International Publishing, 2020, pp. 519–535. DOI: 10.1007/978-3-030-58545-7_30

[2] C. Mirarchi, M. Gholamzadehmir, B. Daniotti, and A. Pavan, "Semantic enrichment of bim: The role of machine learning-based image recognition", *Buildings*, vol. 14, no. 4, p. 1122, Apr. 2024. DOI: 10.3390/buildings14041122

[3] N. Gard, A. Hilsmann, and P. Eisert, "Spvloc: Semantic panoramic viewport matching for 6d camera localization in unseen environments", in *Computer Vision – ECCV 2024*. Springer Nature Switzerland, Dec. 2024, pp. 398–415. DOI: 10.1007/978-3-031-73464-9_24

[4] J. Kim, C. Choi, H. Jang, and Y. M. Kim, "Piccolo: Point cloud-centric omnidirectional localization", in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 3313–3323.

[5] H. Yu, W. Zhen, W. Yang, and S. Scherer, "Line-based camera pose estimation in point cloud of structured environments", *arXiv preprint arXiv:1912.05013*, 2019. DOI: 10.48550/ARXIV.1912.05013

[6] P. Yin, L. Xu, J. Zhang, H. Choset, and S. Scherer, "l3dloc: Image-to-range cross-domain localization robust to inconsistent environmental conditions", *arXiv preprint arXiv:2105.12883*, 2021. DOI: 10.48550/ARXIV.2105.12883

[7] P. Pauwels et al., “A semantic rule checking environment for building performance checking”, *Automation in Construction*, vol. 20, no. 5, pp. 506–518, Aug. 2011. DOI: 10.1016/j.autcon.2010.11.017

[8] T. H. Kolbe, “Representing and exchanging 3d city models with citygml”, in *3D Geo-Information Sciences*. Springer Berlin Heidelberg, 2009, pp. 15–31. DOI: 10.1007/978-3-540-87395-2_2

[9] H. Butler, M. Daly, A. Doyle, S. Gillies, T. Schaub, and S. Hagen, *The GeoJSON Format*, RFC 7946, Aug. 2016. DOI: 10.17487/RFC7946 [Online]. Available: https://www.rfc-editor.org/info/rfc7946

[10] R. B. Rusu and S. Cousins, “3d is here: Point cloud library (pcl)”, in *2011 IEEE International Conference on Robotics and Automation*, 2011, pp. 1–4. DOI: 10.1109/ICRA.2011.5980567

[11] A. Aziz, N. Gard, M. König, P. Eisert, J. H. Heinbach, and L. Trost, “Visual fire safety inspection framework using computer vision algorithms”, *Journal of Computing in Civil Engineering*, vol. 39, no. 5, Sep. 2025. DOI: 10.1061/jccee5.cpeng-6492

[12] A. C. Barreiro, M. Trzeciakiewicz, A. Hilsmann, and P. Eisert, “Automatic reconstruction of semantic 3d models from 2d floor plans”, in *2023 18th International Conference on Machine Vision and Applications (MVA)*, IEEE, 2023, pp. 1–5.

[13] H. Taira et al., “Inloc: Indoor visual localization with dense matching and view synthesis”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7199–7209.

[14] S. Wang, S. Fidler, and R. Urtasun, “Lost shopping! monocular localization in large indoor spaces”, in *2015 IEEE International Conference on Computer Vision (ICCV)*, IEEE, Dec. 2015. DOI: 10.1109/iccv.2015.309

[15] C. Liu, A. G. Schwing, K. Kundu, R. Urtasun, and S. Fidler, “Rent3d: Floor-plan priors for monocular layout estimation”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3413–3421.

[16] A. Soultana and A. Aziz, “Active learning approach for object detection in technical building equipment images”, in *34th Forum Bauinformatik / 34. Forum Bauinformatik (Bochum, 06. - 08.09.2023)*, 2023. DOI: 10.13154/294-10103

Reconstruction of IFC Models Using Voxelization for Spatial Analysis

Cemre Karaoglan[1]
[1]Design Computation, RWTH Aachen University, Germany
E-mail(s): karaoglan@dc.rwth-aachen.de

Abstract: In most large-scale building projects, contractor companies receive finished architectural designs. The IFC models in these packages are the keystones of the construction automation tasks for the following phases; however, they are often riddled with inconsistencies. So, the models are then either cast aside or remodeled from scratch, causing immense inefficiency. To address this issue, this paper proposes an open-source methodology that can restructure the IFC models by rebuilding their geometry without losing the semantic integrity. This is achieved through voxelization - by converting the model into a 3D grid representation. Although voxelization has been explored in BIM-related research, its application for building-scale spatial analysis remains limited. This approach contributes to bridging the gap between design and construction, enabling practitioners to use external IFC models reliably. The methodology is benchmarked against an existing solution with diverse models. Although there is no significant improvement, the results prove the current study to be an alternative approach with potential. Future work includes improving accuracy through volumetric sampling and scalability with multiprocessing.

Keywords: Building Information Modeling (BIM), Voxelization, Rasterization, Geometry Processing

DOI: 10.18154/RWTH-CONV-254886. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Industry Foundation Classes (IFC)[1] is an established industry standard for collaboration between various stakeholders for Building Information Modeling (BIM) models in construction projects. Accurate IFC data are crucial to enable reliable and efficient automation of construction tasks such as quantity take-off, costing, scheduling, and simulation. Many general contractor firms take over the projects after the tendering phase and receive already completed design projects, with IFC models in the package. However, these external IFC models often suffer from inconsistent, broken, and/or missing geometric and semantic data, so it becomes challenging to use them reliably for spatial analysis tasks and construction management workflows. As a result, they are either not used or are modeled from scratch, resulting in a significant waste of time, energy, and money. This clearly shows the need for a tool that can rebuild IFC models to ensure the geometric and semantic fidelity for further automation tasks, especially when companies lack in-house planning experts.

[1]buildingSMART IFC4.0 Documentation

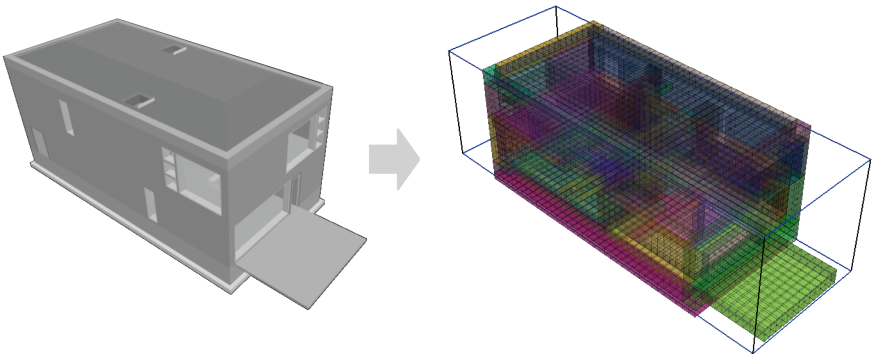


Figure 1: Voxelization of an IFC model - Duplex model example.

With this motivation, a rasterization-based method is developed for reshaping IFC geometries. Rasterization means the conversion of an image into individual pixels that can be displayed on a screen. In this study, the same principle is applied to the IFC model for the conversion of 3D geometries into voxels, thus voxelization (Figure 1). By transforming polygonal geometric descriptions into structured 3D grid representations, the proposed approach can address both geometric robustness and computational tractability. Moreover, the voxel-grid structure holds the potential to enable efficient spatial queries and optimized data storage by assigning the semantic metadata to voxel cells. This way, it may be possible to repair the broken component in the BIM workflow and fill in the missing piece that occurs during the tendering process between the design phase and the construction phase.

2 State of the Art

Voxel grids serve not only as efficient geometric proxies but also as rich data structures for linking semantics, automation, and environmental feedback into BIM workflows. This versatility positions voxelization as a powerful tool for a deeper understanding and analysis of the geometry in the built environment and thus has diverse application areas.

In the context of energy analysis and simulation, voxelization is often used for converting IFC model contents into watertight geometries required for thermal simulation. For example, the open-source Voxelization Toolkit for (IFC) Building Models [1] creates enclosed spaces by filling geometric gaps in imperfect BIM models [2]. Numerous other studies also benefit from voxelization to prepare the geometries for energy analysis [3], [4], [5].

The integration of geospatial modeling and BIM requires down-sampling of the geometries for scalability. Naturally, voxel-based approaches offer robust solutions to geometry simplification operations. For this purpose, Khan et al. [6] use adaptive voxel grids to represent heterogeneous soil data, linking it to BIM and CityGML for urban subsurface modeling. In another study, Khan et al. [7] propose a parametric voxel method for earthwork planning in connection to geotechnical attributes. Due to the scale in geotechnical models, these studies prioritize semantic richness and volumetric classification over more precise geometric representation formats.

A 3D grid structure is inherently easier and computationally less expensive to navigate in comparison to a collection of objects scattered in the modeling space. So, voxelization comes as a valuable

tool in converting the space into a more machine-readable format. Therefore, lately it has been increasingly used in spatial reasoning studies, such as path-finding, evacuation simulations, and semantic segmentation. Xie et al. [8] represent negative space with voxels and classify them based on motion types for evacuation simulations, while Aleksandrov et al. [9] create connectivity graphs from navigable voxel space. Martens and Blankenbach [10] use voxel-based segmentation for indoor partitioning, and Tran et al. [11] and Chen et al. [12] model coverage zones for camera placement. Luttun and Krijnen [13] further apply deep learning to voxelized IFC models, using 3D convolutional networks for automated semantic segmentation of building geometry. Additional applications include geometry repair [14], semantic exterior tagging [15], scan-to-BIM automation [16], and digital twin updates for damage assessment [17].

These studies reflect the diverse use of voxelization in BIM-related domains. As seen in this section, recent studies highlight its potential for advanced tasks such as spatial reasoning and classification. However, almost none of the studies focus specifically on repairing or restructuring IFC models at the building scale for construction automation tasks, e.g., quantity takeoff and collision check. So, this further underlines the need for a method that can support practitioners in repairing external IFC model geometries. In this study, we aim to address this need by creating an opportunity to utilize IFC models without requiring BIM tool intervention. Although its primary focus lies in the energy analysis applications, the voxelization toolkit [1] has the capacity to fulfill this need as well. Therefore, in the following parts, it will be used as a benchmark to measure the performance of the current methodology.

3 Methodology

In this section, one possible rasterization approach to restructuring IFC model geometries is introduced. Detecting the boundaries of each geometry and colliding them with a grid cell in the same space is a costly geometric operation. So, the process is simplified and replaced by a more computationally efficient approach. Specifically, polygon rasterization is used as the core methodology. This is achieved in three steps: First, by converting IFC descriptions into geometric representations, then populating these geometries with points, and lastly by casting the points onto a 3D structured grid (Figure 2).

3.1 Approach

Initially, the process begins with converting IFC geometry descriptions into 3D meshes to further process them. While IFC models are semantically rich and hierarchical, these meshes are geometric and optimized for spatial operations. With the aim of preserving the semantics, each IFC element is assigned an integer value - a *label* so to say. These *labels* are stored in a dictionary with their corresponding IFC element data. This step simplifies the IFC structure, discards redundant information, and prepares the dataset for point cloud generation.

Secondly, meshes are converted into point clouds using uniform surface sampling. To maintain the geometrical details in the model, the number of sampling points N is calculated in correlation to the voxel size. It is computed as:

$$N = \text{int}(\text{surface_area} \times \text{points_per_unit_area} \times \frac{1}{\text{voxel_surface_area}}) \quad (1)$$

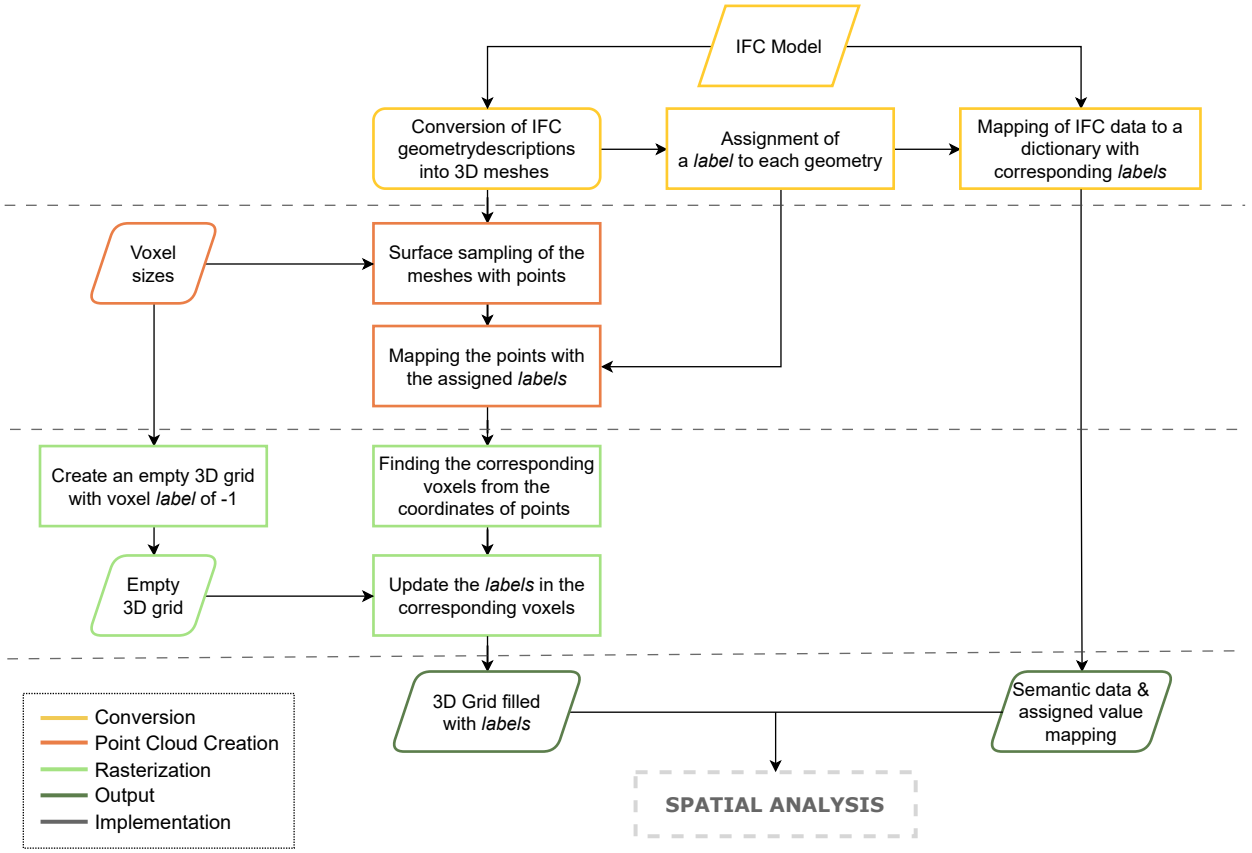


Figure 2: Flowchart of the voxelization process.

Each mesh is individually converted into a point cloud. During this conversion, each point is assigned the *label* of its corresponding IFC element. This way, we keep the link between the semantic data from the IFC model and the point cloud representations.

Lastly, the voxelization step is executed. A structured 3D grid is placed within the outermost boundaries of the model. The size of the voxels is user-defined, and each voxel can be assigned one or more values. Upon creation of the grid, each voxel is initialized with a default value of -1, representing the empty space. As the coordinates x, y, z of each point are already known alongside the grid bounds and voxel sizes, it becomes possible to find their corresponding voxel index, thus their location within the grid.

Points are mapped to cell indices (i, j, k) using the given algorithm:

$$i = \left\lfloor \frac{x - x_{\min}}{\text{voxel_size_x}} \right\rfloor, \quad j = \left\lfloor \frac{y - y_{\min}}{\text{voxel_size_y}} \right\rfloor, \quad k = \left\lfloor \frac{z - z_{\min}}{\text{voxel_size_z}} \right\rfloor \tag{2}$$

Each occupied voxel is assigned the *label* of the point it envelops. where the unoccupied voxels remain with the value of -1. The result is a 3D grid with labeled voxels, ready for spatial queries and analysis (Figure 3).

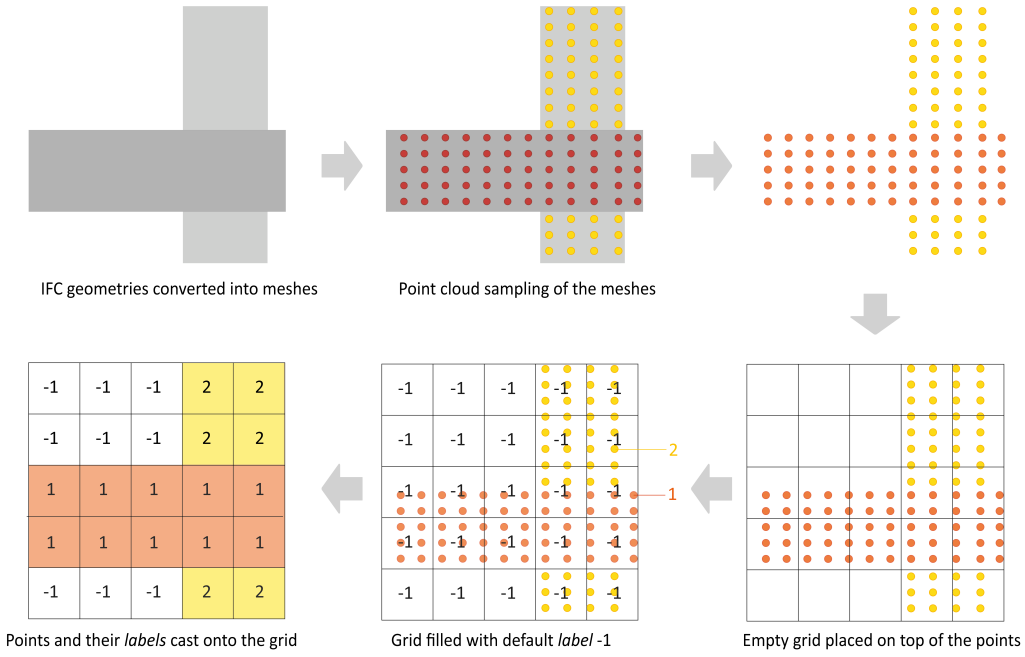


Figure 3: Voxelization of IFC Models - Each drawing is a hypothetical section, and it shows the intersection of a wall (yellow) and a slab (orange).

3.2 Use-Case and Results

The methodology is implemented with Python programming language and Open3D [18], and PyVista [19] libraries. To evaluate performance under real-world conditions, the proposed workflow is tested with seven different models: Research standard Duplex and Clinic models [2] and four models from the industry. The number of polygons in the models are as following: Clinic Model - 63,406; Duplex Model - 9,420; Industry Model A - 16,440; Industry Model B - 39,280; Industry Model C - 93,712; Industry Model D - 67,754.

Each test is applied with a set of point sampling values per unit area (200,300 and 500 pts.) and a set of voxel sizes (0.1, 0.3 and 0.5 m). The existing toolkit [1] is also tested with the same voxel sizes in order to provide benchmark values. The results are presented below in Figure 4.

To test the accuracy of the voxelization process and to see its viability for QTO tasks, IfcSlab and IfcWall volumes are recorded. The actual volumes are acquired with the help of a model checking software and these values are compared to the calculated values from the current study and the existing toolkit. The tests are conducted with 0.1 m voxel size. The accuracy given in the charts represents the calculated/actual volume in percentages (Figure 5).

4 Discussion

The test results given in the prior section have indicated several improvement points. As expected, the processing times grow proportionally to model complexity and grid resolution. Smaller voxels offer higher precision but require more resources, while larger voxels improve performance at the cost of accuracy. Therefore, without further algorithmic improvements, the voxel size must be tuned based

[2] BIM Data Research and Development- Open Source IFC Models

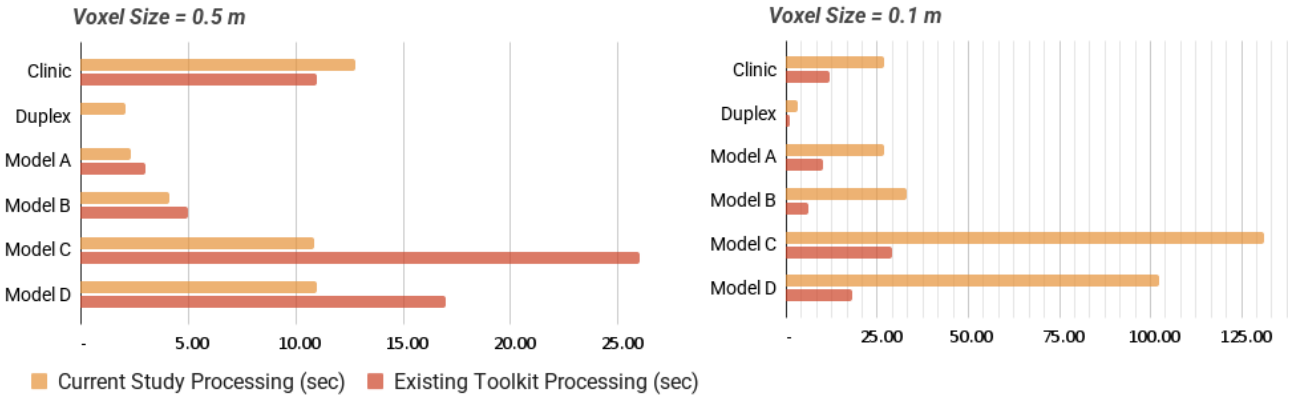


Figure 4: Comparison of processing times in different voxel sizes: Current study vs the existing toolkit [1].

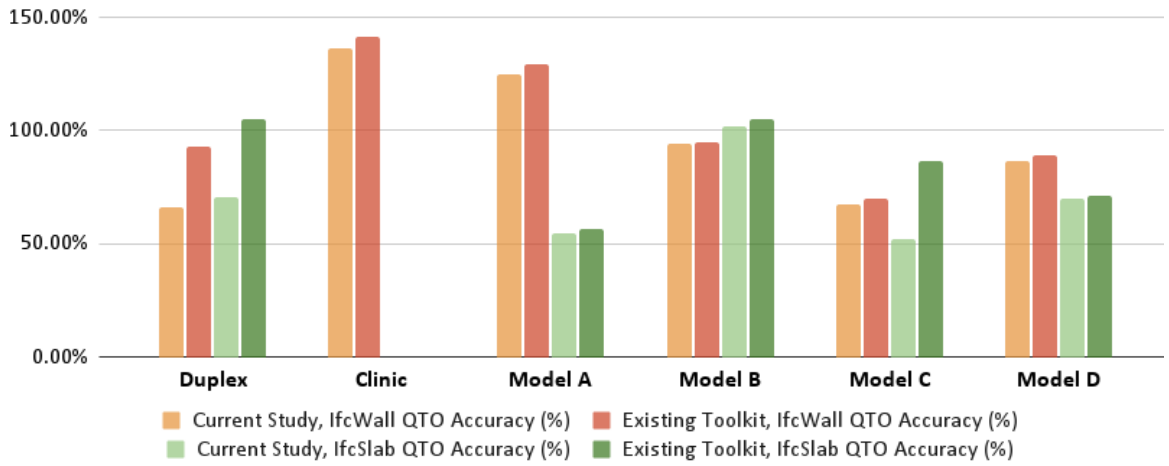


Figure 5: Comparison of quantity take-off accuracies for IfcSlab and IfcWall: Current study vs the existing toolkit [1].

on the use case to maintain an effective trade-off. Although this study is outperforming the existing toolkit in processing times for the 0.5 m voxel size, the toolkit performs much more efficiently when the grid has a higher resolution with 0.1 m voxel size. This shows the need for a more scalable algorithm benefiting from multiprocessing.

For the QTO tests, both models perform similarly, often failing to come close to 100% accuracy due to under- or overestimation. Yet, the existing toolkit takes the lead in most models. Therefore, it is necessary to take measures to improve the algorithm's accuracy when representing the original model. This may be possible through volumetric point sampling rather than surface sampling to represent the total volumes of the model elements. However, volumetric sampling may bring more computational load making scalability even a bigger priority for future work.

Although the results fall below the existing solution and do not fully answer the research questions, even acquiring comparable results is encouraging. The potential range of implementations and applicable future improvements can make this study a strong open-source candidate for restructing external IFC models, filling the gap between design and construction phases.

5 Conclusion

This paper has introduced an open-source methodology for restructuring IFC models through voxelization. By converting polygonal geometries into structured 3D grids while preserving semantic information, the approach addresses the common inconsistencies of external IFC models and prepares them for automation tasks. The benchmarking results demonstrate that, although the current implementation does not yet surpass the existing toolkit in terms of accuracy and scalability, it achieves comparable outcomes and validates the feasibility of the method.

The study underlines the potential of voxel-based approaches to act as a bridge between design data and construction requirements, particularly in workflows where external IFC models are otherwise discarded or remodeled from scratch. Even though the presented workflow is still at a prototypical stage, its open-source nature creates opportunities for further development and adaptation by the community. Future work will focus on improving accuracy through volumetric sampling and enhancing scalability with multiprocessing, but also on testing additional use cases such as negative space creation, connectivity analysis, and model repair. In this way, the methodology may evolve into a robust and versatile solution that supports practitioners in integrating external IFC models more reliably into construction processes.

References

[1] T. Krijnen, K. Arroyo Ohori, and luzpaz, *Voxelization toolkit for (IFC) building models*, Sep. 2023.

[2] T. Krijnen, T. El-Diraby, T. Konomi, and A. Attalla, “Thermal analysis of IFC building models using voxelized geometries”, 2021.

[3] V. Bazjanac, “Space Boundary Requirements For Modeling Of Building Geometry For Energy And Other Performance Simulation”, in *CIB W78: 27th International Conference*, 2010.

[4] C. M. Rose and V. Bazjanac, “An algorithm to generate space boundaries for building energy simulation”, *Engineering with Computers*, vol. 31, no. 2, pp. 271–280, Apr. 2015. DOI: 10.1007/s00366-013-0347-5 Accessed: May 23, 2023.

[5] G. N. Lilis, G. I. Giannakis, and D. V. Rovas, “Automatic generation of second-level space boundary topology from IFC geometry inputs”, *Automation in Construction*, vol. 76, pp. 108–124, Apr. 2017. DOI: 10.1016/j.autcon.2016.08.044 Accessed: May 23, 2023.

[6] M. S. Khan, I. S. Kim, and J. Seo, “A boundary and voxel-based 3D geological data management system leveraging BIM and GIS”, *International Journal of Applied Earth Observation and Geoinformation*, vol. 118, p. 103 277, Apr. 2023. DOI: 10.1016/j.jag.2023.103277 Accessed: Apr. 17, 2023.

[7] M. Khan, J. Kim, S. Park, S. Lee, and J. Seo, “Methodology for Voxel-Based Earthwork Modeling”, *JOURNAL OF CONSTRUCTION ENGINEERING AND MANAGEMENT*, vol. 147, no. 10, Oct. 2021. DOI: 10.1061/(ASCE)CO.1943-7862.0002137

[8] R. Xie, S. Zlatanova, J. Lee, and M. Aleksandrov, “A Motion-Based Conceptual Space Model to Support 3D Evacuation Simulation in Indoor Environments”, *ISPRS International Journal of Geo-Information*, vol. 12, no. 12, 2023. DOI: 10.3390/ijgi12120494

[9] M. Aleksandrov, S. Zlatanova, D. Heslop, and A. Diakite, “BIM-based connectivity graph and voxels classification for pedestrian-hazard interaction”, *Journal of Spatial Science*, vol. 69, no. 3, pp. 699–719, 2024. DOI: 10.1080/14498596.2023.2281923

[10] J. Martens and J. Blankenbach, “VOX2BIM+-A Fast and Robust Approach for Automated Indoor Point Cloud Segmentation and Building Model Generation”, *PFG-JOURNAL OF PHOTOGRAMMETRY REMOTE SENSING AND GEOINFORMATION SCIENCE*, vol. 91, no. 4, pp. 273–294, Aug. 2023. DOI: 10.1007/s41064-023-00243-1

[11] S.-T. Tran, D. Lee, H. Pham, L. Dang, C. Park, and U.-K. Lee, “Leveraging BIM for Enhanced Camera Allocation Planning at Construction Job Sites: A Voxel-Based Site Coverage and Overlapping Analysis”, *Buildings*, vol. 14, no. 6, 2024. DOI: 10.3390/buildings14061880

[12] Z. Chen, Z. Lai, C. Song, X. Zhang, and J. C. Cheng, “Smart camera placement for building surveillance using OpenBIM and an efficient Bi-level optimization approach”, *Journal of Building Engineering*, vol. 77, p. 107 257, Oct. 2023. DOI: 10.1016/j.jobe.2023.107257

[13] J. Luttun and T. Krijnen, “IfcVoxNet: 3D segmentation and classification of voxelized IFC building models with Deep Learning”, 2024.

[14] S. Bischoff, D. Pavic, and L. Kobbelt, “Automatic restoration of polygon models”, *ACM Transactions on Graphics (TOG)*, vol. 24, no. 4, pp. 1332–1352, 2005.

[15] X. Liu, C. He, H. Zhao, J. Jia, and C. Liu, “ExteriorTag: Automatic semantic annotation of BIM building exterior via voxel index analysis”, *IEEE Computer Graphics and Applications*, vol. 41, no. 3, pp. 48–58, 2021. DOI: 10.1109/MCG.2021.3069856

[16] L. Truong-Hong and R. Lindenbergh, “Extracting structural components of concrete buildings from laser scanning point clouds from construction sites”, *Advanced Engineering Informatics*, vol. 51, p. 101 490, Jan. 2022. DOI: 10.1016/j.aei.2021.101490

[17] Y. Gao, H. Li, W. Fu, C. Chai, and T. Su, “Damage volumetric assessment and digital twin synchronization based on LiDAR point clouds”, *Automation in Construction*, vol. 157, p. 105 168, Jan. 2024. DOI: 10.1016/j.autcon.2023.105168

[18] Q.-Y. Zhou, J. Park, and V. Koltun, “Open3D: A modern library for 3D data processing”, *arXiv:1801.09847*, 2018.

[19] B. Sullivan and A. Kaszynski, “PyVista: 3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK)”, *Journal of Open Source Software*, vol. 4, no. 37, p. 1450, May 2019. DOI: 10.21105/joss.01450 [Online]. Available: https://doi.org/10.21105/joss.01450

An Evaluation of Zero-shot Classification of Multi-Modal Foundation Models for Point Clouds

Nils Sören Krause[1], Miguel Guzman-Merino[1] and Dennis Grauer[1]
[1]AI for sustainable construction, University of Rostock, Rostock, Germany
E-mail(s): nils.krause@uni-rostock.de, miguel.merino@uni-rostock.de, dennis.grauer@uni-rostock.de

Abstract: This paper investigates the use of various Multi-Modal Foundation Models (MMFMs) for classifying isometric views of point clouds. It addresses the challenge of reducing the amount of training data required for classification. MMFMs are prompted via their APIs using 2D-views from point cloud data. The paper examines the number of views needed to achieve efficient and accurate object classification in point clouds. We achieve an accuracy of up to 66%, and an F1-score of 35% is reached without specific training on the benchmark dataset. The results provide initial insights into the selection of MMFMs for zero-shot semantic segmentation of point cloud data. The approach is validated on parts of the ModelNet40 benchmark dataset.

Keywords: Multi-Modal Foundation Models, Classification, Zero-shot, Point cloud

DOI: 10.18154/RWTH-CONV-254887. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

As a step towards new labeled datasets in civil engineering, this paper examines the potential of Multi-Modal Foundation Models (MMFMs) in zero-shot classification of point clouds. It focuses on the MMFMs GPT-4o and Gemini-1.5-pro, as well as a reversed Stable Diffusion (Stable Diffusion-v1-5) model, in a comparative analysis. The test data is obtained from the ModelNet40 benchmark dataset [1]. The issue under discussion is whether the models can generate value labels for pre-classifying a dataset. Furthermore, the question arises as to whether the models are helpful for the semantic segmentation of the data. It has been demonstrated that these algorithms are capable of acquiring knowledge in a manner analogous to that of a human, through the utilization of previously classified data [2], [3]. The preparation of this classified data for the machine by a human is a laborious process. To reduce the effort required for semantic segmentation of point clouds, we are testing the behavior of classifications using Multi-Modal Foundation Models.

The paper begins with a State-of-the-art section, followed by the methodology and a demonstration of the tests. Afterwards, the test results and a discussion follow. Finally, we provide an outlook on further areas where research is needed. The results of this study help us to understand how to use MMFMs for point cloud classification.

The main contributions of the paper are:

- Evaluating GPT-4o, Gemini-1.5-pro and Stable Diffusion v1.5 for classifying isometric views of point clouds.
- Testing of different strategies for prompting the MMFM.
- Testing of different numbers of isometric views for classification.
- Classification without pre-trained models.
- Testing with a Benchmark dataset to ensure reproducible results.

2 State-of-the-art

Since the introduction of ChatGPT in 2022, the development and use of MMFMs for various applications have accelerated rapidly. Models such as GPT-4o from OpenAI, Gemini from Google, Claude from Anthropic, or DeepSeek demonstrate that they can process text, images, videos, and code. While they have shown potential in generative tasks–such as creating architectural images [4], floor plans [5], or assisting in the building design phase [6]–the understanding of point clouds remains underexplored. The most significant challenge is analyzing the unstructured data derived from LiDAR scans, for which deep learning techniques are employed [7]. There are still differences in the methods to classify point clouds; some approaches use raw point clouds [8], [9] for the segmentation and classification of point clouds. However, these approaches can only recognize data with which they are familiar. The performance of these models depends heavily on specific parameters that must be adapted to the models. They also require extensive datasets for training or testing of these models, as demonstrated by [10]. There are still datasets for point clouds [11], [12]. However, it leaks labeled datasets, particularly in civil engineering. To obtain these datasets for civil engineering, we need to examine how to label the point cloud data with minimal manual effort. For the point cloud pre-processing task, the State-of-the-art is also looking for new direction, they are going to use large language models for the point cloud alignment [13]. Based on these, we will take a deeper look at Prompt variation and the capabilities of MMFMs for image recognition of point clouds. To use MMFMs for point cloud understanding, we transform the point cloud data into images by generating isometric views. The works of [14] and [15] enable us to classify point clouds in the form of images, also named zero-shot classification. For example, the work of [16] presents a model approach named PointBLIP for zero-shot classification of point clouds, tested on ModelNet40, with an accuracy of 66.25 %.

3 Methodical

The 3D-data originates from 3D-mesh objects collected as OFF files in the ModelNet40 dataset [1]. In step (i), the ModelNet40 dataset was downloaded. The ModelNet40 dataset consists of 40 classes that can be utilized for semantic segmentation. In step (ii), the OFF objects contained within the data set were converted into point clouds. The point clouds consist of area-weighted distributions with exactly 3000 points. In step (iii), one to four isometric views were extracted at the following angles: *front, side, top, and diagonal*. Images contain a white background and black dots, sized at 512×512 pixels, exported in PNG-format. In step (iv), these views were then prompted into the various MMFMs via an API connection. In step (v), ensure its alignment with the established ground-truth for effective

classification based on ModelNet40 labels. The entire code and used labels for data preparation and evaluation can be viewed in our GitHub [1]. In Figure 1, you can see an example of the point cloud from the used dataset ModelNet40. We did not use the complete dataset because the API costs for prompting were too high.

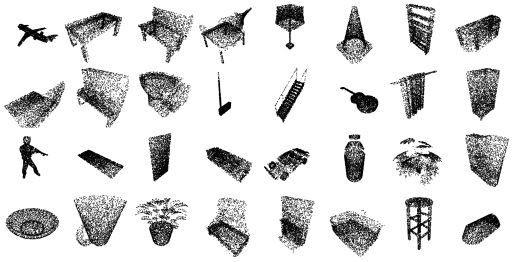


Figure 1: Isometric views of the point clouds from the used ModelNet40 dataset.

3.1 Multi-Modal Models

For the evaluation, several current MMFMs were selected. Based on their ability to process visual and textual input via their APIs, priority was given to models with public availability. Stable Diffusion was chosen due to its captioning capability, allowing for a comparison of the visual understanding of different models. The following models in Table 1 were selected:

Table 1: Overview of tested MMFMs.

Model	From	API	Multi-Modal skills
GPT-4o	OpenAI	Yes	Image + Text
Gemini-1.5-pro	Google	Yes	Image + Text
Stable Diffusion	Stable Diffusion-v1-5	local	Image + Text(encoder)

3.2 How to test

The following prompt in Algorithm 1 was used for evaluation. The prompt only varies in the words: *"three isometric"* for the respective number of isometric views:

Algorithm 1: Prompt for isometric classification.

```
prompt = (
You are shown three isometric views of a 3D object. Choose exactly one of the following
    words that best describes the
object based on the three isometric views. Reply with only one word from this list (no
    extra text):
join(sorted(list(class_list)))
```

The ground-truth data is based on the labels of ModelNet40. The individual classes were transferred to the MMFMs as a class list, which is shown in Algorithm 2, so that they do not deviate from the ground truth.

Algorithm 2: Class list for the prompt.

[1] https://github.com/AI4SC/Multimodal-point-cloud-classifier.git

```
class_list = {
"airplane", "bed", "bench", "bookshelf", "bottle", "bowl", "car", "chair", "cone", "cup",
    "curtain", "door", "flower_pot",
"glass_box", "guitar", "keyboard", "lamp", "laptop", "mantel", "person", "piano", "plant"
    , "radio", "range_hood",
"sink", "stairs", "stool", "tent", "toilet", "tv_stand", "vase", "wardrobe", "xbox"}
```

The evaluation is then carried out using the following metrics:

- **Accuracy**: The accuracy is the overall performance of the true predicted labels by the different models. The accuracy is expressed as a percentage of true predicted labels.
- **F1-Score**: Averages the F1-score for each class without considering class imbalances, known as the macro-average F1-Score.

3.3 Real data testing

We tested the workflow using synthetic data and then applied it to real data scanned with the BLK2GO from Leica. The data was segmented from the point cloud using CloudCompare. We generated a single labeled point cloud for each object in a class. The scene contains a 3D-printer, beamer, desktop, cardboard box windows, heater, lamp, shelf, robot arm, robot dog, chair, and table. The key difference lies in our ability to extract geometric features from the point cloud, including height, brightness, width, and color. The geometric features are derived from an axis-aligned bounding box (AABB) using Opend3D with Python. The scene is shown in Figure 2. The dataset for testing was modified to ensure that the objects have their real dimensions. For the evaluation of the influence of geometric features, the analysis is conducted, including consideration of these features. Due to the effort required for segmentation, only 35 objects that are present in the scene were tested.

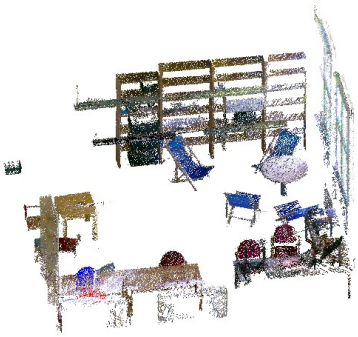


Figure 2: Tested scene of the scanned real data, captured with the BLK2GO

The class list in Algorithm 2 was updated to reflect the new objects. The prompt for testing the models is shown in Algorithm 3. The objects were tested based on four isometric views with the different models.

Algorithm 3: Prompt for isometric classification with AABB.

```
prompt = (
```

```
You are shown four isometric views of a 3D object. The object has the following geometric
    properties
(note: the height, width, and depth are derived from the objects axis-aligned bounding
    box, and may
not represent precise physical dimensions but give a general sense of size and
    proportions)
Height: features['height']
Width: features['width']
Depth: features['depth']
Based on the isometric views and the geometric dimensions, choose exactly one word from
    the list below that best
describes the object. Respond with only one word from this list: join(sorted(class_list))
```

4 Results

The evaluation of the data is presented in Figure 3 and Table 2, which includes the model, samples, accuracy, and F1-score. We evaluated the classification capability of the model GPT-4o and Gemini-1.5-pro, as well as a reversed Stable Diffusion model with a captioning function. The results show that the MMFM Gemini-1.5-pro has better results in every validation scenario. Overall, the classification was even more precise with colored point clouds and the addition of geometric features. The results indicate that the captioning function in Stable Diffusion can be used to generate a partial description of the classification of one isometric view of the point clouds. Filters were used for captioning processing, which not only searched for the required class label but also for synonyms similar to the class. The model achieved an accuracy of 17 %.

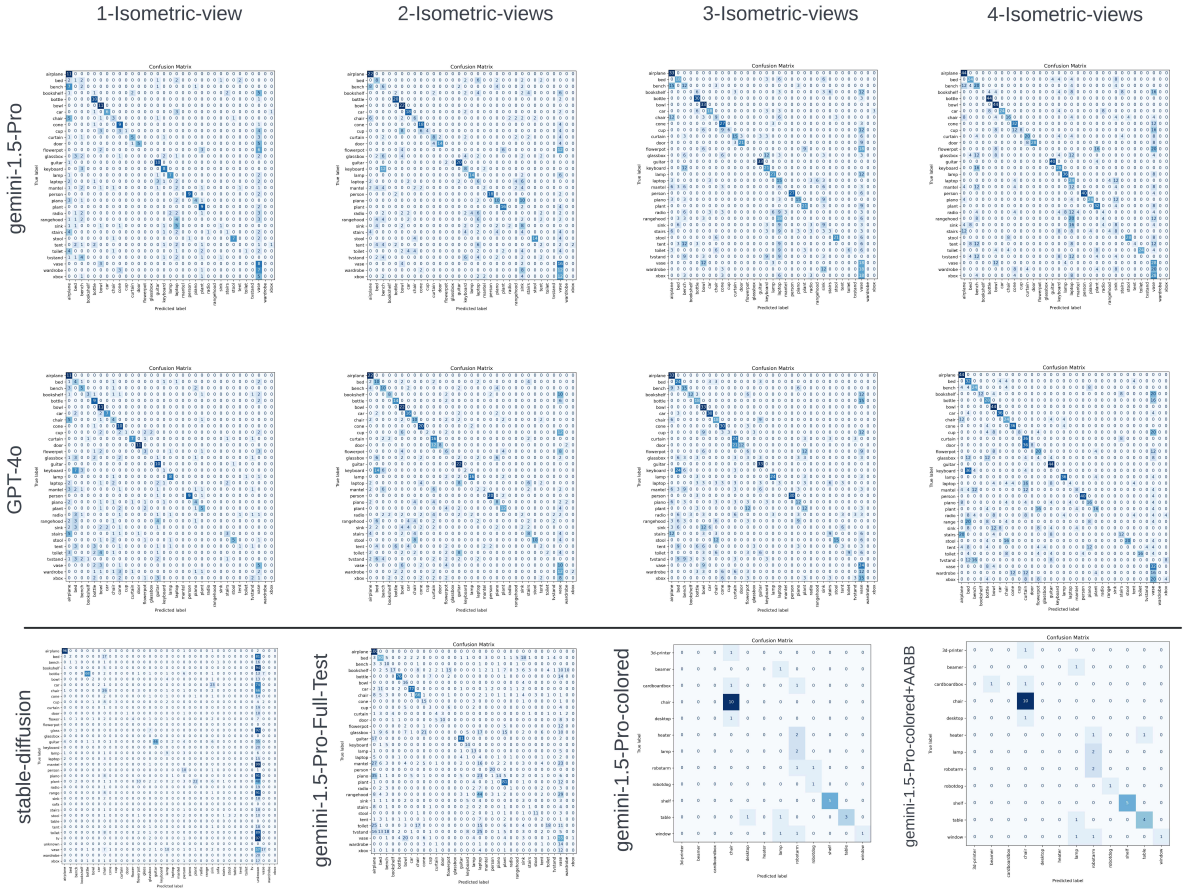


Figure 3: Confusion Matrix's of the classification results, from the evaluated models(First row: Gemini-1.5-pro, 1-4 Isometric views; Second row: GPT-4o, 1-4 Isometric views; Third row: Stable Diffusion: ModelNet40, test data big test, Gemini-1.5-pro-big-test: ModelNet40, big test, Gemini-1.5-pro-colored: colored real point cloud BLK2GO, Gemini-1.5-pro-colored: colored real point cloud BLK2GO + AABB).

Table 2: Classification accuracy and F1-score of GPT-4o and Gemini-1.5-pro on isometric views and different evaluations.

Model Variant	**Samples**	**Accuracy**	**F1-score**
GPT-4o-1-Iso.-view	363	0.38	0.32
GPT-4o-2-Iso.-views	726	0.39	0.33
GPT-4o-3-Iso.-views	1089	0.40	0.34
GPT-4o-4-Iso.-views	1452	0.41	0.35
Gemini-1.5-pro-1-Iso.-view	363	0.36	0.32
Gemini-1.5-pro-2-Iso.-views	726	0.39	0.35
Gemini-1.5-pro-3-Iso.-views	1089	0.40	0.36
Gemini-1.5-pro-4-Iso.-views	1452	0.43	0.39
Gemini-1.5-pro-1-Iso.-view-big-test	1861	0.41	0.32
Stable-Diffusion-1-Iso.-view; CLIP/ViT-L-14/openai-big-test	1861	0.17	0.14
GPT-4o-4-Iso.-views-colored-cloud	35	0.54	0.29
GPT-4o-4-Iso.-views-feature-geometries	35	0.57	0.29
Gemini-1.5-pro-4-Iso.-views-colored-cloud	35	0.60	0.33
Gemini-1.5-pro-4-Iso.-views-feature-geometries	35	0.66	0.35

5 Discussion and outlook

The evaluation results presented in Table 2 and Figure 3 provide insights into the classification capabilities of two MMFMs (GPT-4o and Gemini-1.5-pro). In initial tests, the performance of the classification improves with the increasing number of isometric views provided. For example, the GPT-4o model increases from 38 % to an accuracy of 41 %. In contrast, the Gemini-1.5-pro model increased the accuracy from 36 % to 43 %. These results confirm that the classification with multi-view input of point clouds has a positive influence. For Future development, we approach to test the use of more than four isometric views. Beyond the visual input, we tested the benefits of enhancing the prompt with additional features from the point cloud. Classification performance increases if the point cloud is colored. Otherwise, the implementation of geometric features from AABB as textual input to the model demonstrates an understanding of models for context information of objects. For example, the accuracy in classification for tables increases with the known possible geometric dimension of the object. But these need to be evaluated in further studies. The Stable Diffusion model achieves an accuracy of 17 % in captioning images. Stable Diffusion accuracy is 25 % lower than the MMFM results from the Gemini-1.5-pro test. A closer look at the results from the confusion matrix in Figure 3 shows differences in the classification results of the objects. We can observe that, for example, the GPT-4o model struggles to recognize the following objects: glass boxes, keyboards, radios, range hoods, sinks, tents, TV stands, and wardrobes. The Gemini-1.5-pro model has issues with the following objects: bookshelves, flowerpots, glass boxes, mantels, sinks, stairs, TV stands, wardrobes, and Xboxes. But the Gemini-1.5-pro model was able to recognize these objects with a higher hit rate, whereas GPT-4o had problems with radios, tents, and keyboards. Both models had difficulties with these objects: glass boxes, sinks, TV stands, and wardrobes. Now, look at the geometric properties in Figure 1. For example, whether it is a rectangular object or a circular object, it becomes opaque. The objects that the model has recognized incorrectly are only partly similar in their geometric properties. We can recognize this clearly with the confusion matrix, which is based on the predicted labels and the actual class of the vase. Both models accurately labeled a wide range of objects in that column, but they struggled with mostly rectangular structures. Below the diagonal of the confusion matrix, Gemini shows a slight cluster of misclassified labels between the airplane and flowerpot classes. In contrast, GPT-4o distribution is spread across the airplane, up to the bench and laptop classes. Furthermore, objects such as planes, guitars, people, and chairs are easily recognized by the models, as they stand out clearly from other objects that usually consist only of a rectangular or circular surface. This suggests the model lacks by details to recognize the objects. From the distribution of the other objects in the confusion matrix. The Stable Diffusion model shows that it is rather intended for the generation of labels for images of point clouds. For future work, we tested the semantic segmentation for construction site objects. Figure 4 shows the results of the outlook.

ANSWER: This object appears to be a tripod stand, most likely a:
Construction laser tripod or **support stand.**
Key visual clues:
Three legs splayed at the base – standard tripod design. Vertical central pole with an adjustable height mechanism. Possibly a mounting clamp or laser bracket near the top.

ANSWER: The object in this image appears to be a metal scaffolding structure, commonly used in construction or building maintenance. Key visual indicators:
The vertical and diagonal metal poles **form a typical scaffolding frame**. There's a working platform or plank at the top. The structure is freestanding and modular, designed to support workers or materials temporarily. There is some protective railing and possibly safety equipment visible.

ANSWER: The image appears to show a 3D point cloud or photogrammetric reconstruction of a construction site object, specifically a stack of building materials.
Most likely interpretation:
The large central object looks like: **A pallet or stack of wooden planks or panels**, possibly drywall sheets, wood panels, or insulated panels, commonly found on construction sites. On top of the stack are some long wooden beams or planks.

Figure 4: Outlook of GPT-4o-Model with point cloud data from a construction site. The figure shows the answers of the model after prompting the model with the question: What is it?

6 Conclusion

The results show that available MMFMs are alternative tools for the pre-classification of new datasets. These datasets are important because well-segmented and semantically segmented point cloud data are not widely available in the construction sector. This work represents a step toward new methods and capabilities for obtaining labeled point cloud data more easily and quickly by using MMFMs. But they still have limitations in terms of accuracy or in the prompting chain of the model. The approach can benefit from chain-of-thought reasoning and expanding the geometric information, e.g., with density histograms or eigenvalues based on shape ratios. But this is a point for future research. The comparison of the MMFMs used with other pre-trained deep learning approaches for point cloud classification on ModelNet40 does not come close to these, in the range of 66 % to 99 % [16] [2]. Our evaluation reached a maximum of 43 % with ModelNet40 and 66 % with real data.

References

[1] Z. Wu et al., "3d shapenets: A deep representation for volumetric shapes", in *IEEE conference on computer vision and pattern recognition*, 2015.

[2] H. Zhang et al., "Deep learning-based 3d point cloud classification: A systematic survey and outlook", *Displays*, 2023.

[2] https://modelnet.cs.princeton.edu/

[3] E. Grilli, F. Menna, and F. Remondino, “A review of point clouds segmentation and classification algorithms”, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2017.

[4] V. Paananen, J. Oppenlaender, and A. Visuri, “Using text-to-image generation for architectural design ideation”, *Journal of Architectural Computing*, 2024.

[5] J. Ploennigs and M. Berger, “Automating computational design with generative ai”, *Civil Engineering Design*, 2024.

[6] C. Zhong, L. Yi’an Shi, and L. Wang, “Ai-enhanced performative building design optimisation and exploration”, in *29 Conference on Computer-Aided Architectural Design Research*, 2024.

[7] S. A. Bello, S. Yu, C. Wang, J. M. Adam, and J. Li, “Deep learning on 3d point clouds”, *Remote Sensing*, 2020.

[8] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, “Pointnet: Deep learning on point sets for 3d classification and segmentation”, in *IEEE conference on computer vision and pattern recognition*, 2017.

[9] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, “Pointnet++: Deep hierarchical feature learning on point sets in a metric space”, *Advances in neural information processing systems*, 2017.

[10] G. Qian et al., “Pointnext: Revisiting pointnet++ with improved training and scaling strategies”, *Advances in neural information processing systems*, 2022.

[11] M. De Deuge, A. Quadros, C. Hung, and B. Douillard, “Unsupervised feature learning for classification of outdoor 3d scans”, in *Australasian conference on robitics and automation*, 2013.

[12] M. A. Uy, Q.-H. Pham, B.-S. Hua, T. Nguyen, and S.-K. Yeung, “Revisiting point cloud classification: A new benchmark dataset and classification model on real-world data”, in *CVF international conference on computer vision*, 2019.

[13] Z. Guo et al., “Point-bind & point-llm: Aligning point cloud with multi-modality for 3d understanding, generation, and instruction following”, *arXiv:2309.00615*, 2023.

[14] N. Zhao, T.-S. Chua, and G. H. Lee, “Few-shot 3d point cloud semantic segmentation”, in *IEEE/CVF conference on computer vision and pattern recognition*, 2021.

[15] R. Zhang et al., “Pointclip: Point cloud understanding by clip”, in *IEEE/CVF conference on computer vision and pattern recognition*, 2022.

[16] Y. Xiao, Y. Dou, and S. Yang, “Pointblip: Zero-training point cloud classification network based on blip-2 model”, *Remote Sensing*, 2024.

Evaluating the DIN EN ISO 11855 1D-RC-approach for modeling temperature profiles of pipe embedding façade systems

Oliver Abele[1], Hannah Riefer[1], Jérôme Frisch[1] and Christoph van Treeck[1]
[1]Institute of Energy Efficiency and Sustainable Building (E3D), RWTH Aachen University, Aachen, Germany
E-mail(s): abele@rwth-aachen.de, hannah.riefer@rwth-aachen.de, frisch@rwth-aachen.de, treeck@rwth-aachen.de

Abstract: New façade technologies that play an active role within a building's energy system are gaining in relevance. This paper investigates a clinker brick wall with an embedded pipe system to collect solar heat gains. DIN EN ISO 11855-2 provides a modeling approach to translate underfloor heating systems into resistance capacitance models. While this approach can model the heating power of underfloor heating systems, its suitability to assess internal temperature profiles of clinker walls remains unknown. This paper aims to close this gap by simulating the given system both in the 1D approach using Matlab Simulink as well as in a 2D finite element approach using WUFI 2D. The 2D simulation allows to evaluate temperature differences within the wall that the 1D approach is incapable of modeling. Three different model setups are investigated both in 1D and 2D. The results show that the approach investigated is suitable to simulate the heat gains of the system. The inclusion of air perforations in the models led to improvements in modeling the heat gains but also to stronger divergence between 1D and 2D temperature profiles, where the 1D models overestimated the depth of the impact of the pipe.

Keywords: DIN EN ISO 11855-2, clinker facade, WUFI, Matlab, pipe embedding facade systems

DOI: 10.18154/RWTH-CONV-254888. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

To combat the global challenge of climate change, new façade technologies that can actively save or even produce energy are being developed. In particular, this paper investigates pipe embedding façade systems that allow for heat flows to be directed by means of a fluid flowing through these pipes. The advantage of such systems is that this turns the façade from a passive barrier that separates the inside from the outside environment into an active component of the building's energy system that can be adapted to differing boundary conditions.

In the existing literature, pipe embedding façade systems are used for different purposes and applications. Multiple studies explore embedded pipes as a way of creating a switchable insulation that can

intercept heat flow between inside and outside when desired [1], [2]. In [3], the solar energy on a roof is simultaneously collected for domestic hot water generation.

The focus of the existing literature on similar systems lies mostly on a system level, analyzing how the system can interact with the whole building's energy system and influence the building's energy consumption. The simulation approaches applied are mostly specific to the respective use case and often include a co-simulation of multiple models [2], [4], making them difficult to adapt for other projects.

To reduce the computational complexity of simulations, allowing them to be used for different application fields such as advanced control strategies, a simplification of models into one dimensional resistance capacitance (RC) models is desirable. To achieve this, [5] uses 2D numerical simulations in COMSOL to calibrate the model's resistances. DIN EN ISO 11855-2 [6] provides a standardized simulation approach to model the heat transfer of underfloor heating systems to the adjacent rooms by introducing a fictive layer to the RC model. In [7], this approach has been used in model predictive control (MPC). Façade systems differ from the norm's applicability domain of underfloor heatings in various aspects such as façades being exposed to outside weather conditions, differing building materials (the norm applies only to pipes embedded in either wood-constructions, concrete or screed) and purpose of the respective systems.

Therefore, this paper aims to investigate the norm's applicability to a pipe embedding clinker brick façade that collects solar heat gains of the wall to be used as a source for a heat pump. Due to their air perforations as well as mortar layers in between bricks, clinker façades are heterogeneous even without the inclusion of embedded pipe systems, making them challenging to translate into 1D models. The first aim of this paper therefore is to evaluate the 1D model's accuracy in calculating the total heat gathered by the system. For this purpose, three different simulation model setups, each gradually introducing more details, are investigated. By removing heat from an outside wall, the given system may interfere with the façade's moisture balance either by causing condensation within the wall or by hindering the wall's drying phases which might lead to moisture induced damage. To assess this risk of condensation, detailed temperature profiles inside the walls are needed. The second aim of this paper is to examine the accuracy of the resulting temperature profiles by comparing them to the results of 2D simulations carried out in WUFI 2D.

2 Methodology

2.1 Experimental Setup

The experimental data used to validate the simulation results was gathered over six days in January 2018. The experiments were conducted on a test façade similar to the one presented in [8]. It consists of a wall of energy clinker brick measuring 8 m². In total, 90.64 m of pipe are embedded in the wall, distributed into 47 rows of clinker. The setup only features a layer of clinker, so no other layers that would be present in a real building (insulation, load bearing brick wall) are considered in this study. Both sides of the wall were exposed to the outdoor environment with the side embedding the pipe facing south. Measurements of the volume flow, the inlet and return temperatures of the fluid and the solar radiation on the wall surface were available. Outdoor temperature and wind speed were not measured

directly and therefore had to be extracted from historical weather data from Germany's National Meteorological Service (DWD) for the closest weather station (Münster / Osnabrück, Germany) [9].

2.2 Simulation Models

This study compares simulations in 1D and 2D for three different simulation setups, resulting in a total of six simulation models, as depicted in Figure 1. In the following, first the respective modeling approaches for both 1D and 2D simulations are explained. Then, the three different model setups and their specific differences are shown.

The 1D simulations are carried out in Matlab / Simulink using the Simscape Library. Following the DIN EN ISO 11855-2 approach, the heat exchange between the pipe and the wall is modeled by introducing a virtual heat conducting layer to the model that is connected to the temperature within the pipe via an additional virtual resistance R_x. In the wall's direction (left to right in Figure 1), the models are discretized into eleven 2R1C blocks, allowing for a detailed overview of internal temperatures within the brick. Alongside the pipe's axis, the model is discretized into 47 equal elements, each one representing one row (1.9 m) of clinker bricks.

The 2D simulations are conducted in WUFI 2D. Preliminary investigations have shown that the temperature loss in one row does not significantly impact the rows below or above. Therefore, a model of three rows of clinker with adiabatic boundary conditions at the top and bottom of the model representing additional clinker rows was chosen. Since WUFI 2D cannot simulate the flowing liquid and its convective heat exchange with the wall, the pipes are instead modeled as heat sinks. The data for these heat sinks is taken from the 1D Matlab simulation because the experimental data only contains the total heat gain of the whole wall instead of the necessary detailed heat gains from individual rows. The topmost three rows were chosen because this is where the fluid enters the wall and the highest heat losses occur. The pipe was modeled as a square with the same surface area as its real life circular counterpart. Since this study only focuses on thermal aspects, water content and humidity within the materials were disregarded.

All models use the same material data, which was taken from the clinker's product declaration when available and approximated using standard material databases otherwise. To settle on plausible starting conditions, the simulations include an adjustment period of roughly three days of stationary conditions before the start of the six days of experiment.

The 1D and 2D models are compared for three different model setups which are depicted in Figure 1. These model setups get progressively more detailed and therefore introduce more and more differences between the one and two-dimensional simulations.

- Model setup "Solid" models the clinker as a solid homogeneous mass, ignoring the clinker's air perforations. This approach represents the most common and simplest way of modeling clinker bricks in 1D simulations. The differences between the two models are the pipe, which is modeled explicitly in the 2D model while being translated to a virtual layer in the 1D model and the mortar layer in between bricks, which is ignored in the 1D model.

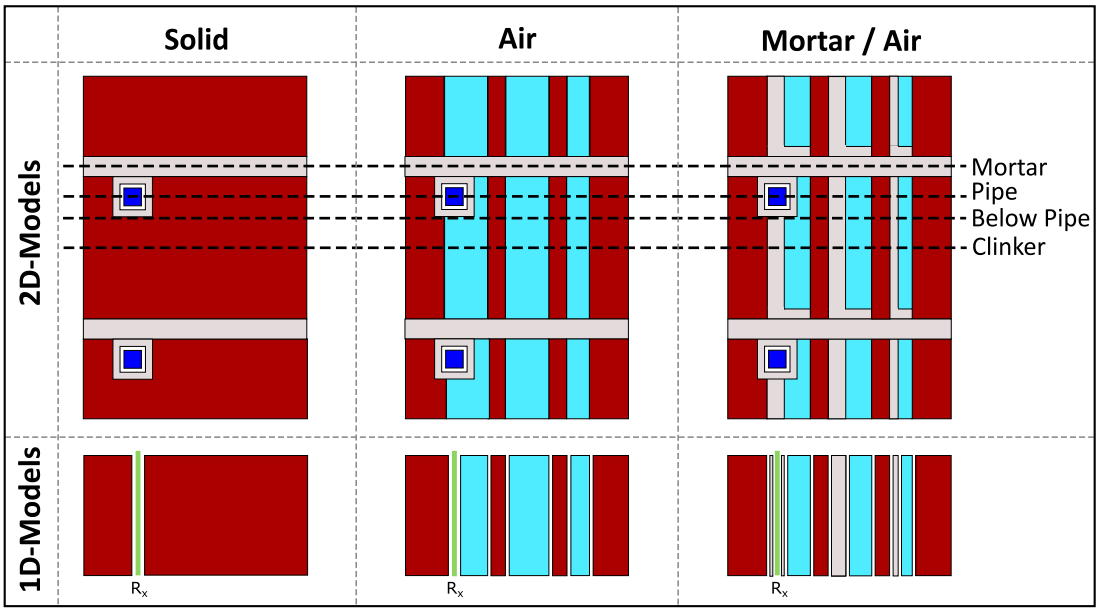


Figure 1: Simulation Model setups and evaluation heights

- Model setup "Air" introduces the air perforations within the clinker brick to both models. While this does not add an additional difference between the models, this setup increases the importance of the mortar in between bricks, which can act as a thermal bridge between the solid bricks.
- Model setup "Mortar / Air" adds partial mortar filling to the perforations. This mortar can enter the perforations either by falling down from the top of the brick when applying the mortar for the next brick layer or from the bottom when the brick is pressed onto the mortar bed during installation. Another paper currently in press demonstrated how taking this excess mortar into account can influence simulation results.

To evaluate the simulations, first the return temperature of the fluid (after leaving the energy clinker façade) is compared between the 1D simulations and the experimental data. This comparison highlights how the different setups influence the overall performance of using 1D simulations to simulate the total heat gains of the system. Second, the inner temperature profiles within the brick are compared between the 1D and 2D models. While the 1D simulations only result in one temperature curve for the complete height, the profiles of the 2D simulations are examined at four different evaluation heights (Mortar, Pipe, Below Pipe, Clinker) as depicted in Figure 1. The evaluation is conducted for the topmost clinker brick, where the heat losses are the highest. The root mean squared error (RMSE) and coefficient of determination (R^2) are used to quantify the goodness of fit for all comparisons.

3 Results

3.1 Generated Heating Power

Overall, all three 1D models manage to capture the general trends of the fluid return temperature as depicted in Figure 2. The air model reaches the best accuracy of the three with an RMSE of 1.6 °C and an R^2 of 0.88. Differences between the models are most noticeable during the days where

the solid model (R^2: 0.82) mostly underestimates the return temperature and the air / mortar model (R^2: 0.85) overestimates the return temperature. During the nighttime, all three models show very similar results, all three underestimating the measured return temperature. A possible reason for the discrepancy between day- and nighttime could be the radiation data (which plays the most important role in heating up the wall during the day) being measured directly and therefore more accurately than the ambient air temperature which was taken from a weather station.

The results show that the DIN EN ISO 11855-2 approach is suitable to simulate the heat exchange between fluid and façade even for heterogeneous building materials such as clinker brick. Including an air layer to represent the brick's perforations does however influence the results, most likely because it acts as an insulation layer within the brick, thus "trapping" the heat within the frontmost part of the brick where it can be exchanged with the fluid.

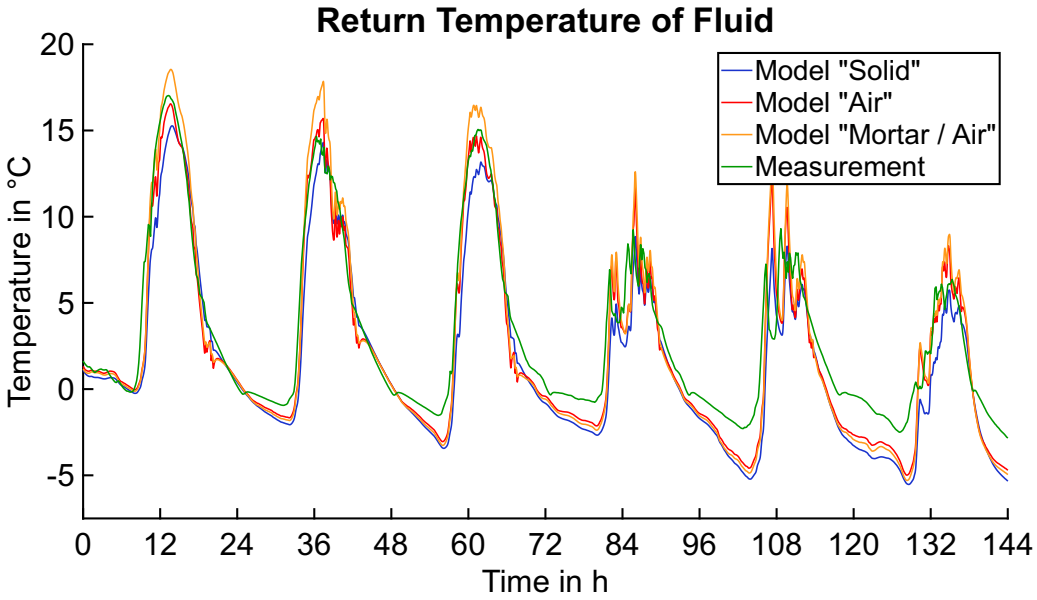


Figure 2: Simulated Return Temperatures of the fluid

3.2 Comparison of 1D and 2D Simulations

Figure 3 compares the temperature curves within the brick of the 1D and 2D simulations for two different points during the experiment: at the temperature peak of the first day of the experiment and at the lowest point of the fourth night of the experiment. During the nighttime, the temperature of the fluid was lower than the surrounding brick's, meaning that the system heated up the façade during that time. It should be noted that the temperature at the height of the pipe (orange line) around the pipe depth does not represent the temperature of the fluid correctly and should therefore be taken with some caution.

Comparing the 2D temperature curves of the three models, the most notable difference can be seen at the sun facing surface, where the air model reaches higher temperatures than the solid and mortar / air models due to its lower thermal mass before the first perforation. At the depth of the pipe, the air model also has the biggest divergence between the different heigths, with the clinker height being much less affected by the heat loss than in the other two models. Towards the rear of the clinker

behind the first perforation, temperature curves for all three models converge, with the solid model showing the overall smoothest and most even distribution due to having no obstructions.

The accordance between 1D and 2D simulations is highest for the solid model. At the depth of the pipe the 1D temperature lies between the temperature just below the pipe and the temperature at mortar level. At further depths, all temperatures, including the 1D temperature, converge to very similar levels. In the models that include the air perforations, the 1D and 2D simulations show similar agreement at the pipe depth but start diverging behind the first perforation. This divergence is most pronounced towards the the end of the third perforation and becomes smaller towards the rear surface.

The RMSE values between 1D and 2D for the whole span of the experiment (Figure 4) underline these findings. While for the solid model, the RMSE gradually decreases behind the pipe, the RMSE for both models including perforations increase towards the end of the third perforation only dropping again at the non-sun-exposed surface. This increase of RMSE for both models is mainly caused by 1D temperatures being lower than all temperatures of the 2D model. This means that the 1D model overestimates until which depth the heat loss through the pipe impacts temperatures.

Overall, the RMSE of the mortar / air model is highest for most depths and height levels. There are two reasons for this: First, the mortar / air model had the biggest heat exchange between wall and fluid (as shown earlier). This means that this model also features the biggest local influence of the pipe. Second, the 2D model includes fully connected mortar layers at the bottom of the perforation, increasing the area of mortar that can act as a thermal bridge, creating more divergence between the 1D and 2D models.

4 Conclusions and Limitations

This study investigated how well the DIN EN ISO 11855 approach of modeling underfloor heating systems in 1D can be employed to simulate heat gains and inner temperature curves of a pipe embedding clinker façade system. Three models were investigated, a solid model only consisting of clinker, a model that adds air perforations within the brick, and a model that adds air perforations as well as some mortar inside these perforations.

The results show that the approach is suitable to model the heat gains of the described system for the given experiment data. Models that incorporate the brick's air perforations calculated higher heat gains of the fluid due to the air acting as an insulation that hinders rapid distribution of heat through the brick. This also led to a slightly better agreement between experimental and simulation data for these models.

To evaluate the resulting temperature curves within the brick, 2D simulations were conducted in WUFI and compared to the 1D results of the DIN approach. Overall, the solid model showed the highest agreement between 1D and 2D models while the model incorporating both air and mortar had the highest differences. Discrepancy between 1D and 2D models especially arose after the first perforation behind the pipe for both models including air layers where the temperatures calculated by the 1D approach were lower than those calculated by the 2D models. This overestimation of the pipe's influence towards the rear end of the brick should be accounted for when trying to employ 1D simulations to determine temperatures within a brick façade.

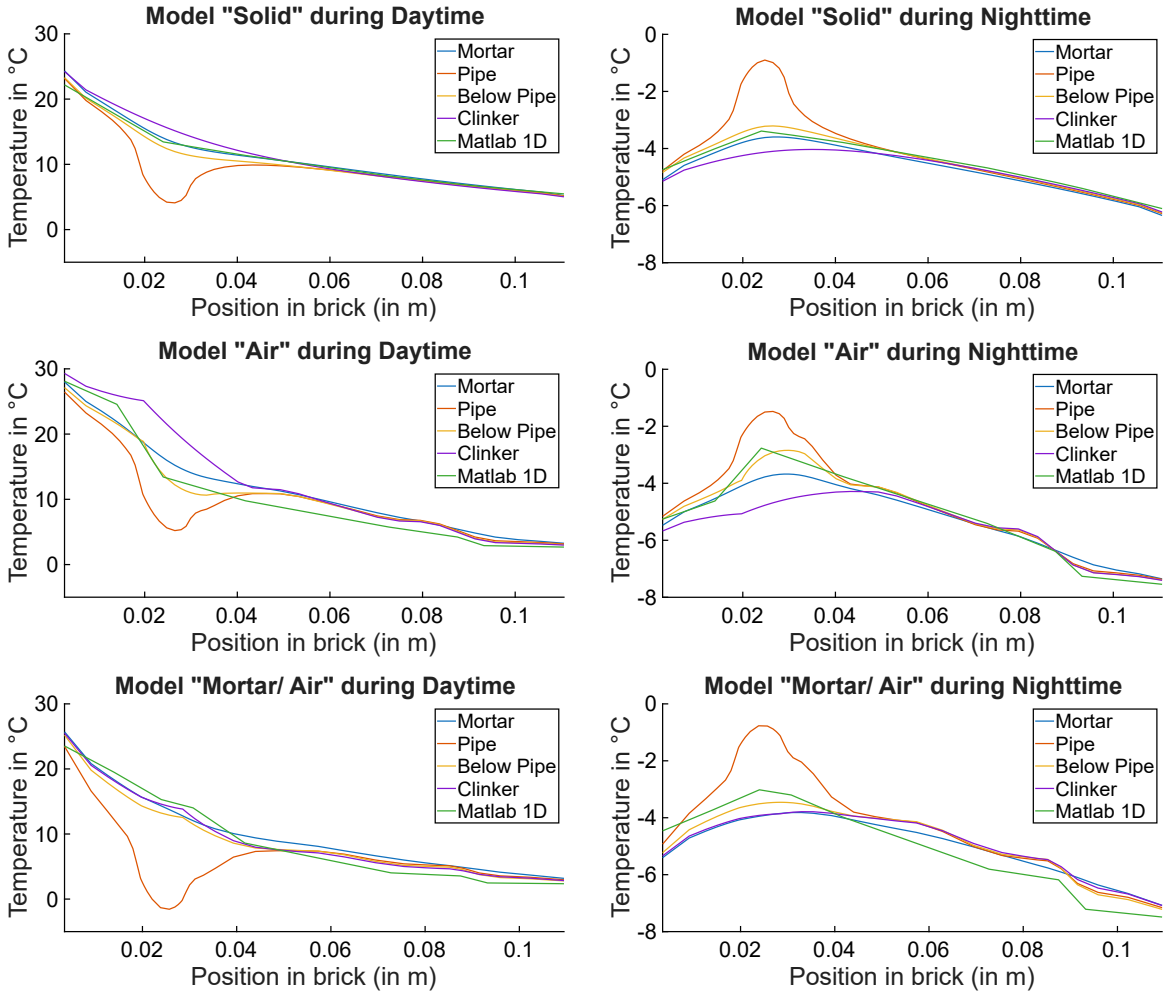


Figure 3: Temperature Profiles of 1D and 2D Simulations for Day and Nighttime

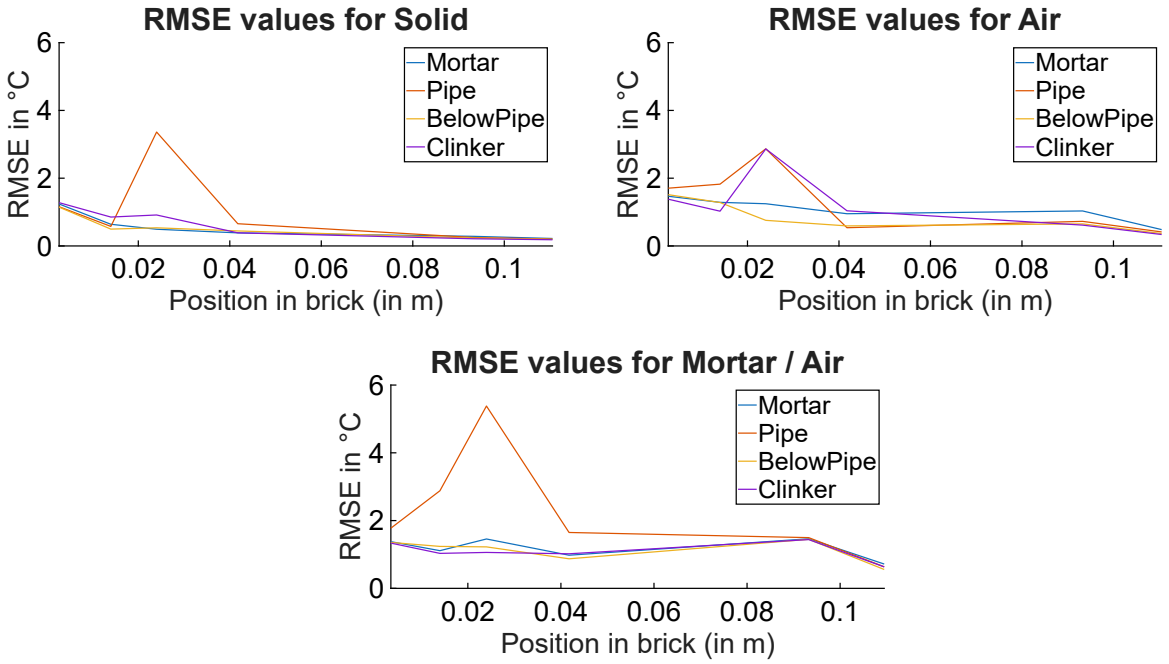


Figure 4: RMSE between 1D and 2D Simulations calculated over the span of the whole experiment

The findings highlight the potential of the DIN EN ISO 11855-2's approach even for heterogeneous materials such as clinker bricks and give some indications for further research. They are however based only on a small experiment of six days with some uncertainty around ambient conditions. The experiment did not monitor inner temperatures of the brick, meaning that the comparison of temperature curves between 1D and 2D simulations could not be validated using real data. Further research should therefore conduct more experiments equipped to measure temperatures at different depths and for different boundary conditions.

Acknowledgements

The authors gratefully acknowledge the financial support of the German Federal Ministry for Economic Affairs and Energy within the project "ReFaTEk" (promotional reference number 03EN6018A).

References

[1] F. U. Vogdt and M. Helbach, "Eine thermisch aktiv gedämmte, monolithische Außenwand aus Infraleichtbeton als nachhaltiges, zukunftsfähiges Wandsystem?", *Bauphysik*, vol. 37, no. 1, pp. 38–51, Feb. 2015. DOI: 10.1002/bapi.201510005 Accessed: Jun. 3, 2025.

[2] Q. Zhu, X. Xu, J. Gao, and F. Xiao, "A semi-dynamic model of active pipe-embedded building envelope for thermal performance evaluation", *International Journal of Thermal Sciences*, vol. 88, pp. 170–179, Feb. 2015. DOI: 10.1016/j.ijthermalsci.2014.09.014 Accessed: Jun. 3, 2025.

[3] R. Sarachitti, C. Chotetanorm, C. Lertsatitthanakorn, and M. Rungsiyopas, "Thermal performance analysis and economic evaluation of roof-integrated solar concrete collector", *Energy and Buildings*, vol. 43, no. 6, pp. 1403–1408, Jun. 2011. DOI: 10.1016/j.enbuild.2011.01.020 Accessed: Jun. 3, 2025.

[4] M. Ibrahim, E. Wurtz, P. H. Biwole, and P. Achard, "Transferring the south solar energy to the north facade through embedded water pipes", *Energy*, vol. 78, pp. 834–845, Dec. 2014. DOI: 10.1016/j.energy.2014.10.078 Accessed: Jun. 3, 2025.

[5] M. Hermann et al., "TABSOLAR - Solarabsorber und andere thermisch aktive Bauteile aus Ultrahochleistungsbeton (UHPC)", Schlussbericht, 2015.

[6] *DIN EN ISO 11855-2:2022-04, Umweltgerechte Gebäudeplanung_- Flächenintegrierte Strahlungsheiz- und -kühlsysteme_- Teil_2: Bestimmung der Auslegungs-Heiz- und Kühlleistung (ISO_11855-2:2021); Deutsche Fassung EN_ISO_11855-2:2021*. DOI: 10.31030/3275482 Accessed: Jul. 4, 2025.

[7] T. H. Dokka, D. N. Lassen, T. Johnsen, and H. Koppang, "Validation of a Simplified Model Predictive Control of a Low Exergy Embedded Heating and Cooling System", Jun. 2023. Accessed: May 27, 2025.

[8] D. Mähner, J. Lengers, and C. Brand, ""Energieklinker" – System zur Nutzung solarer Wärmeenergie in Klinkerfassaden", *Bauphysik*, vol. 38, no. 2, pp. 81–87, 2016. DOI: 10.1002/bapi.201610009 Accessed: Jun. 25, 2025.

[9] D. Wetterdienst, *10-minütige Stationsmessungen der Strahlung und Sonnenscheindauer für Deutschland, Version v24.03*, Mar. 2024.

Benchmarking Large Language Models of BIM: Key Requirements for Domain-Specific Evaluation Datasets

Abdelrahman Abouhadid[1] and Clemens Kujat[1]
[1]University of Rostock, Rostock, Germany
E-mail(s): abdelrahman.hadid@outlook.com, clemens.kujat@uni-rostock.de

Abstract: The application of Large Language Models (LLMs) to document-centric workflows in Building Information Modelling (BIM) presents many opportunities to increase productivity through Retrieval-Augmented Generation (RAG). The use cases present unique challenges that render the usage of established benchmarks for evaluating LLM performance in this setting not easily transferable. This paper investigates in which aspects standard LLM evaluation frameworks fall short in assessing performance within the BIM domain and identifies the characteristics of a high-quality benchmark tailored for this purpose. We trace the limitations of benchmarks to the capability of LLMs in capturing the domain-specific complexities of BIM, such as the need for multi-disciplinary knowledge, handling of structured and unstructured data, and the ability to extract and synthesize interconnected information across documents. In our experimental analysis, we apply this benchmark to evaluate state-of-the-art models using a RAG approach. As this is an initial single-case study based on a small dataset, which contains only the modality text, the findings should be seen as an exploratory step. The results show that the models are capable of processing BIM-related textual data effectively via RAG. This highlights the need for investigations on multimodal BIM data, as they may introduce greater challenges for Building Foundation Models.

Keywords: Building information modeling (BIM), Large Language Model (LLM), Benchmark, Retrieval Augmented Generation (RAG)

DOI: 10.18154/RWTH-CONV-254889. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

The Architecture, Engineering, and Construction (AEC) industry has witnessed significant technological advancements with the introduction of Building Information Modeling (BIM). BIM represents a paradigm shift from traditional 2D design methods to a collaborative, data-rich approach that integrates multiple dimensions of information throughout a building's lifecycle [1]. This growing amount of data also needs to be managed, searched and summarized resulting in new challenges.

Concurrently, the field of artificial intelligence has experienced progress with the development of Large Language Models (LLM). These LLMs are designed to summarize and generate human language

across complex domains [2]. Particularly, the introduction of Retrieval Augmented Generation (RAG) approaches [3] allows to inject contextual knowledge from document collections into the LLMs.

However, the integration of LLMs into BIM workflows is not without challenges. The complexity of BIM data structures and the multi-disciplinary character of AEC projects create unique requirements that generic benchmarks fail to address adequately [4], [5]. Current evaluation frameworks for LLMs primarily focus on general language understanding, factual knowledge retrieval, simple reasoning, and single-document comprehension. These frameworks, while valuable in their respective contexts, do not capture the requirements of BIM applications.

This paper seeks to bridge this gap by investigating the limitations of current LLM-based RAG evaluation frameworks in the BIM context and proposing key requirements for domain-specific evaluation datasets. To achieve that, we give an overview of LLMs, RAGs, and benchmarks in section 2. Based on this review we derive a benchmark methodology for civil engineering in section 3. In section 4, we identify critical shortcomings in existing benchmarks and outline essential characteristics that a gold-standard BIM-specific benchmark should possess, and then present the prototype benchmark in section 5. The initial implementation is limited to the text modality and a narrow document set, but it serves as a useful entry point for developing more comprehensive, multimodal benchmarks for Building Foundation Models. Section 6 shows the evaluation of the models, and in section 7, we summarize the small-scale, exploratory study and provide directions for future research.

The significance of this research lies in the following points:

- Identifying key requirements for BIM-specific LLM benchmarks
- Designing and applying a domain-specific benchmark to evaluate LLM performance using a Retrieval-Augmented Generation approach in a text-only setting
- Demonstrating that classical RAG methods can effectively handle Bim-related textual data, motivating further research on the challenges of multimodal BIM data

2 State of the Art

Large Language Models

(LLMs) are generative artificial intelligence models based on the transformer architecture [6]. They are capable of capturing the context of every element in an input sequence (e. g. text) and make predictions for the subsequent element. To achieve this, they are trained on vast corpora of text data, e. g. from the web. Although the training data is very comprehensive, there may be gaps in certain domains, which can lead to incorrect responses, known as hallucinations.

Retrieval-Augmented Generation

(RAG) enables models to first retrieve relevant information from available documents before generating the answer, thereby enhancing the model's knowledge [3]. To achieve that, the content of the documents is divided into chunks, which are embedded and stored in a database, so that they can be retrieved. The chunking is done to keep the context of the model short. When the model is invoked, it

retrieves relevant information from provided documents based on the query using a similarity search, and then, that information gets added to the context of the model.

Benchmarks for LLMs

evaluate the capabilities of LLMs in different use cases. While recent benchmarks such as L-Eval [7], LongBench [8], RULER [9], HELMET [10], Loong [11], and NIAH [12] have advanced the evaluation of long-context LLMs, their general-purpose nature limits their applicability in technical domains like construction and BIM, where capabilities like domain-specific precision, structured data interpretation, cross-document synthesis, and RAG-informed responses are required. Although some benchmarks support multi-document handling, none evaluate model performance on structured data such as tables of asset data, understanding role-based workflows, or validating adherence to technical documentation like BIM Execution Plans or Asset Information Requirements. Moreover, none systematically assesses RAG, despite its importance in grounding BIM queries on domain-specific documentation.

LLM-as-a-judge

has emerged as a methodology for evaluating LLM outputs, wherein one LLM assesses the quality, accuracy, and relevance of another's responses. The approach is more capable of adjusting to the probabilistic answer nature of LLMs than traditional pattern-matching-based benchmark approaches [13]. This offers flexibility in defining evaluation criteria and can be tailored to specific domains. However, its effectiveness depends heavily on the quality of evaluation prompts and the capabilities of the judging model.

3 Methodology

Figure 1 illustrates the methodology followed by the presented study. It splits into a dataset creation phase, where we define a golden benchmark dataset, and an evaluation phase to validate it.

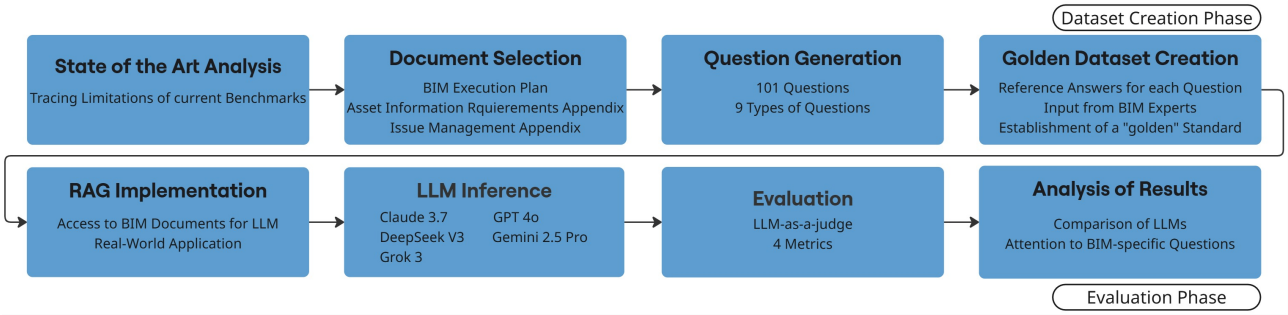


Figure 1: Process of creating and evaluating the domain-specific dataset for assessing LLM performance in BIM-related tasks through a RAG approach.

We first determined the limitations of existing benchmarks with regard to BIM data and then selected relevant BIM documents as the foundation for the evaluation dataset: a BIM Execution Plan (BEP), an Asset Information Requirements Appendix, and an Issue Management Appendix. These documents were chosen as they represent core documentation in BIM implementation, containing a mix of structured and unstructured information that spans technical, managerial, and operational aspects. In the next step, 101 questions were formulated based on our knowledge, covering various aspects of BIM implementation, information management, and technical requirements. The questions were

categorized into eight distinct types designed to test different cognitive abilities and address the identified limitations. Following that, the golden dataset was created, where for each question, reference answers were developed with input from BIM experts to establish a "golden" standard against which LLM responses could be evaluated.

To enable the LLMs to access and process BIM documents when answering questions, a RAG system was implemented to extract the knowledge from the BIM documents and enrich the model's context. This approach was selected over a model tuning, as it allows to quickly add new documents as it is common in real-world applications. Different LLMs were tested with the benchmark questions and their answers were assessed using the "LLM-as-a-judge" methodology [13]. Finally, the performance data was analyzed to identify patterns, strengths, and weaknesses in how different LLMs handled BIM-specific questions, with particular attention to areas where traditional evaluation frameworks might fail to capture important aspects of performance.

4 Requirements for BIM-Specific LLM Benchmarks

The analysis in section 2 revealed several limitations in applying standard benchmarks to the BIM domain, which stem from the unique characteristics of BIM information and workflows that are not adequately captured. Based on our findings, we propose requirements for BIM-specific benchmarks.

(R1) Domain-Specific Knowledge Assessment:

A BIM-specific evaluation dataset must thoroughly assess domain knowledge, including technical terminology, standards, and classification systems (e. g., Uniclass, OmniClass). It should also encompass multi-disciplinary concepts covering knowledge from various disciplines involved in BIM, as well as regulatory and compliance knowledge.

(R2) Evaluation of Structured and Unstructured Data Handling:

The evaluation dataset should assess the ability to work with both structured and unstructured information. This includes the ability to extract and interpret information presented in tables, lists, or other structured document formats, and the assessment of how well models can connect structured data with unstructured contextual information.

(R3) Assessment of Cross-Document Information Synthesis:

The dataset must evaluate the capacity for working across multiple documents and information sources. Therefore, it needs to include information distributed across different documents and sections, and assess the capabilities to identify contradictions or inconsistencies between different information sources to maintain data integrity.

(R4) Evaluation of Context-Aware Reasoning:

There is a need for scenario-based questions, where problems are set within specific project scenarios, that require consideration of multiple contextual factors, mimicking the complexity of real-world BIM applications. Apart from that, numerical reasoning should be assessed to measure the LLMs' ability to perform calculations, quantitative analysis, and mathematical operations.

(R5) Represent BIM Workflow Understanding:

Procedural knowledge should be assessed, using questions about a sequence of activities, information exchanges, and decision points in BIM workflows, testing understanding of how information evolves. Since BIM processes are role-based, an evaluation of understanding regarding responsibilities and interactions between different stakeholders in the BIM context should happen.

(R6) Hallucination Prevention:

In engineering context, it is important that retrieved information is not just accurate, but, also that missing information is identified and communicated back to the user.

(R7) Factual Answers:

In extend to those requirements we also need to ensure that the retrieval of facts is correct. This requirement is not specific for civil engineering, but, of high relevance.

5 BIM-Specific Evaluation Benchmark

5.1 Golden Dataset and Questions

Following the defined requirements for a benchmark, we propose a golden dataset with 101 questions for a real-world BIM dataset. The three BIM documents are analyzed in PDF format, as the proportion of PDF files in the BIM context is up to 40% [14]. The questions are categorized into eight types: (C1) Cross-section reasoning; (C2) List/Table extraction; (C3) Multimodal understanding; (C4) Recognition of not mentioned information; (C5) Numerical reasoning; (C6) General reasoning; (C7) Single-span fact retrieval, and (C8) Yes/No questions. Table 1 presents selected examples from the dataset, illustrating how each question type maps to the requirements from section 4.

Table 1: Representative questions and their targeted requirements

Question Type	Example Question	Targeted Req.
C1	Which software product named in Table 36 is referenced in the mitigation action for “Data inconsistency” in Table 03, and how often is that action scheduled?	R3
C2	List every software platform named in Table 36.	R2
C3	What are the title block instructions mentioned in Figure 14: Project Titleblock (A0)?	R2
C4	What is the 4D Synchro version that the consultant must use?	R6
C5	What percentage of the total design window from Concept Design 100% to Detailed Design 100% is taken up by the Schematic Design phase (50% $\rightarrow$ 100%)?	R4
C6	A drawing that is already in the Published / Issued zone needs revision. In which CDE status zone must the updated file be placed first for review?	R1, R5
C7	What is the planned completion date for Schematic Design 100%?	R7
C8	Must Priority-A (high-severity) clashes be resolved before Tender issue?	R7

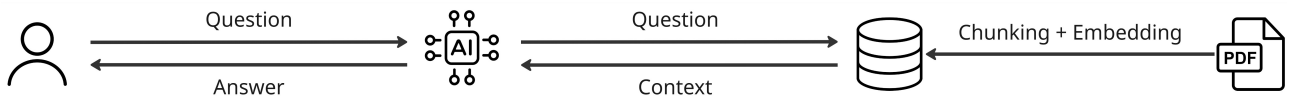


Figure 2: Overview of RAG process: Model receives question, retrieves context from the vector database and responds based on question and context

Algorithm 1: Part of the prompt used as model input

1. *Faithfulness (vs Context-Free)*: How well the answer adheres ***to*** the retrieved context without fabricating information outside of it in the scale 1-10: 10 points: Answer is completely based on the given context; 5 points: Answer is partially based on context with some external information; 1 points: Answer ignores context or fabricates information

2. *Answer Relevance (vs Incomplete)*: How relevant and complete the answer is ***to*** the question in the scale 1-10: 10 points: Answer is completely relevant and complete; 5 points: Answer is relevant but incomplete or contains unnecessary information; 1 points: Answer is irrelevant or very incomplete

3. *Context Relevance (vs Noisy Context)*: How well the most relevant parts of the context are used in the scale 1-10: 10 points: Used the most relevant parts of the context; 5 points: Used some relevant parts ***while*** ignoring important sections; 1 points: Did not use the relevant parts of the context

4. *Correctness*: How accurate the answer is compared ***to*** the golden/reference answer in the scale 1-10: 10 points: Answer is completely correct and aligns perfectly with the golden answer; 5 points: Answer is partially correct with some key information matching the golden answer; 1 points: Answer is incorrect or contradicts the golden answer

5.2 Evaluation Criteria

Given the set of questions for the reference dataset, we need to establish an evaluation methodology for the correctness of their answers. We evaluate the RAG-LLM performance based on the following criteria [3]. The Faithfulness assesses whether the model's response is supported by and consistent with the provided context to prevent hallucinations in RAG systems. The Answer Relevance evaluates how well the generated response addresses the specific query. The Context Relevance measures how relevant the retrieved information is to the query, evaluating the RAG system's ability to extract only the context, which is needed to answer the question. Figure 2 provides an overview of the single steps involved in the RAG, including the context retrieval.

Using the LLM-as-a-judge approach, the judge evaluates the response on all criteria from 1 to 10, given the question and its category, the answer, the reference answer, and the retrieved context. The part of the prompt describing the metrics is shown in Algorithm 1.

6 Experiment

6.1 Setup

The experimental setup includes the selected models and the implementation of the RAG approach. We ran our evaluation on the models Claude 3.7, DeepSeek V3, Grok 3, ChatGPT 4o, and Gemini 2.5 Pro. The RAG documents were processed by chunking their text content with a size of 500 and an overlap of 100, and embedding those chunks with the OpenAI text-embedding-3-large model. The embeddings were then stored in a Pinecone vector database and retrieved by a similarity search and top-3 retrieval. GPT-4o was used as the LLM-judge.

Table 2: Average values of every metric grouped by model

Model	Faithfulness	Answer Relevance	Context Relevance	Correctness
Claude 3.7	9.19	7.70	8.35	6.45
DeepSeek V3	9.28	7.68	8.85	6.19
Gemini 2.5 Pro	8.65	6.81	7.53	6.23
Grok3	9.39	7.89	8.64	6.66
GPT-4o	9.29	6.95	8.61	6.21

Table 3: Average correctness values of every category grouped by model

Model	Cross-section reasoning	List/Table extraction	Multimodal single fact retrieval	Not mentioned	Numerical reasoning	Reasoning	Single-span fact retrieval	Yes/No
Claude 3.7	6.09	6.54	9.00	9.00	4.79	4.23	9.00	5.46
DeepSeek V3	5.64	6.54	9.00	7.75	4.21	4.77	7.00	6.77
Gemini 2.5 Pro	5.27	7.04	8.20	9.44	4.86	3.31	8.25	5.08
Grok3	5.45	6.67	9.00	9.00	5.71	4.23	7.83	7.54
GPT-4o	4.36	5.75	9.00	10.00	4.50	4.08	7.58	7.62
Average	**5.36**	**6.51**	**8.84**	**9.04**	**4.81**	**4.12**	**7.93**	**6.49**

6.2 Results

As the study has an exploratory character, we only ran the benchmark once per model. Table 2 shows that the answers are based on the retrieved context and use most of its relevant context. The Answer Relevance and Correctness are comparably low, meaning that at least for some of the categories, the models' answers do not adequately address the question or give the wrong answer. The average correctness values for each category in table 3 show that the models tend to be able to retrieve single facts out of the provided documents and are also able to detect if the requested information is missing and not start hallucinating. As soon as the models have to access structured information like tables or have to reason over multiple sections, the answers become worse.

6.3 Discussion

The fact that models can extract single facts means that the RAG approach works well with the modality text when there is not much context needed to answer the question.

Some of the questions from the dataset are very specific to the exemplary documents used. For example, the question with category C2 in table 1 would assume that the user knows the document, which undermines the purpose of using an LLM.

The good results in the multimodal category might lead to the assumption that the text-based RAG works well for images. In our case, the specific questions only require the model to retrieve single keywords and don't need to consider the additional information contained in the image.

The results regarding the extraction of structured information, like tables, show that the relationships in the data are lost due to the chunking by the models and that they do not understand tables contextually.

The selected models are not reasoning models, leading to the results in the different reasoning categories. The reasoning category describes a setting where the models have to use implicit context

or domain-specific knowledge not explicitly stated in the documents, which are abilities that are not fully addressed by the selected general-purpose models.

GPT-4o being the only judge might lead to an evaluation bias, for example, it can sometimes be overly strict by penalizing minor differences such as slight variations in datetime formats, and its judgment is less controllable and interpretable compared to explicit metrics.

7 Conclusion and Future Work

This exploratory study developed an initial benchmark for evaluating LLM performance in a RAG approach for BIM data management. The results show that the RAG itself works with BIM documents and also that the models have limitations for some question categories as reasoning or processing of structured data like tables. However, the study is not very representative due to the small set of questions, providing future research directions, such as creating a more generalizable set of questions with a larger size, including the introduction of questions addressing more detailed information in images, to cover the multimodal nature of BIM data. Therefore, a more sophisticated RAG could be used to process images using Building Foundation Models. To assess the quality of the questions from the reasoning categories better, reasoning models could be benchmarked and compared.

References

[1] S. Azhar, “Building information modeling (BIM): Trends, benefits, risks, and challenges for the AEC industry”, *Leadership and management in engineering*, vol. 11, no. 3, pp. 241–252, 2011.

[2] B. Mann et al., “Language models are few-shot learners”, *arXiv preprint arXiv:2005.14165*, vol. 1, p. 3, 2020.

[3] Y. Gao et al., “Retrieval-augmented generation for large language models: A survey”, *arXiv preprint arXiv:2312.10997*, vol. 2, no. 1, 2023.

[4] G. Lee, S. Jang, and S. Hyun, “A generalized LLM-augmented BIM framework: Application to a speech-to-BIM system”, *arXiv preprint arXiv:2409.18345*, 2024.

[5] J. Ploennigs et al., “Building foundation models - potentials, challenges and research directions for using LLM and LVM in AEC”, in *EC3*, 2025.

[6] A. Vaswani et al., “Attention is all you need”, *Advances in neural information processing systems*, vol. 30, 2017.

[7] C. An et al., “L-eval: Instituting standardized evaluation for long context language models”, *arXiv preprint arXiv:2307.11088*, 2023.

[8] Y. Bai et al., “Longbench: A bilingual, multitask benchmark for long context understanding”, *arXiv preprint arXiv:2308.14508*, 2023.

[9] C.-P. Hsieh et al., “Ruler: What’s the real context size of your long-context language models?”, *URL https://arxiv. org/abs/2404.06654*,

[10] H. Yen et al., “Helmet: How to evaluate long-context language models effectively and thoroughly”, *arXiv preprint arXiv:2410.02694*, 2024.

[11] M. Wang et al., “Leave no document behind: Benchmarking long-context LLMs with extended multi-doc QA”, *arXiv preprint arXiv:2406.17419*, 2024.

[12] G. Kamradt, *LLMTest needle in a haystack-pressure testing LLMs*, 2023.

[13] L. Zheng et al., “Judging LLM-as-a-judge with MT-bench and chatbot arena”, *Advances in Neural Information Processing Systems*, vol. 36, pp. 46 595–46 623, 2023.

[14] A. Bach, T. Al-Wesabi, and I. Staka, “Datenzentrierte KI als basis für ein zukünftiges informationsmanagement”, in *Künstliche Intelligenz im Bauwesen: Grundlagen und Anwendungsfälle*, Springer, 2024, pp. 25–39.

Improving Data Robustness in Environmental Monitoring: Lessons from a Green Façade Study

Harald Lesan[1]
[1]Chair of Building Technology (GBT), RWTH Aachen University, Aachen, Germany
E-mail(s): lesan@gbt.arch.rwth-aachen.de.

Abstract: Green façades may reduce building surface temperatures, but quantitative field data remains limited due to the lack of continuous thermal imaging applications. We developed a low-cost experimental setup combining standard temperature sensors with thermal cameras to measure thermal effects of a DIY green façade (BGK) compared to a reference wall over 127 days. Our approach integrated thermal cameras, PT100 probes, black bulb thermometers, and air temperature sensors with local weather monitoring through a custom Python-based workflow. The green façade consistently maintained air temperatures 2K cooler than the reference, with surface temperature differences most pronounced in the evening (-3K to -5K). Radiant temperature measurements revealed differences up to -12K on the hottest day, with minimal differences under cloudy conditions. Results demonstrate both the green façade's cooling potential and the value of continuous thermal imaging for detailed thermal behavior characterization of green infrastructure systems.

DOI: 10.18154/RWTH-CONV-254898. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Recent bibliographic analysis reveals an increase in publications on green façade thermal performance while highlighting a lack of thermal imaging [1] for descriptive thermal behavior analysis. Thermal imaging with IoT devices enables continuous measurements that provide detailed insight into the thermal dynamics of the green façade [2]. However, existing applications are typically limited to short-term studies due to cost and complexity constraints. This study addresses these limitations through collaboration with Landscape Architecture (LA) to monitor their DIY green façade (BGK) system [3]. We developed an experimental setup combining standard sensors with thermal imaging to quantify green façade thermal effects while establishing methodological approaches for long-term thermal behavior characterization through continuous imaging. Our approach demonstrates that low-cost equipment (under 2000€) can capture continuous 5-minute thermal measurements over extended periods, potentially spanning entire annual cycles. This combination of high spatial and temporal resolution provides unprecedented detail for thermal analysis of green infrastructure systems.

2 Materials and Methods

2.1 Site Description and Setup

We selected two adjacent third-floor terraces at the Sprachenzentrum (Eilfschornsteinstraße 15, 52062 Aachen, Germany), each measuring 3.5 m × 8.5 m and bounded by three walls with 50% southwest exposure. The reference terrace maintained its original plastered brick wall, while the experimental terrace hosted the BGK system (2 m × 2 m, 30 cm depth). The BGK features an aluminum frame with adjustable rods for varied substrates and plants, black geotextile fleece pockets for planting, and irrigation via pump and tank with seasonal valve adjustment (see Fig. 9). An adjacent deciduous tree, initially leafless during winter setup, later created uneven shading across measurement periods. Figure 8 shows hourly simulated shading differences.

2.2 Hardware and Configuration

Tinkerforge devices [4] captured thermal and weather data. Figure 2 shows our station types:

Wall Observer (WO): `WO_g` observed the green façade, `WO_r` the reference wall. Each used a thermal camera, a black-bulb PTC sensor, a sound pressure sensor, and motion detection (ultimately yielding no reliable occupant data).

Outdoor Stations (OS): measured air temperature/humidity at ∼1.2 m and on the façade or wall surfaces. `OS_WOg`, `OS_WOr` monitored near observers; `OS_Wg`, `OS_Wr` monitored surfaces.

PTC Stations: PT100 sensors on wall surfaces, see 9. `PTC_WgPo`, `PTC_WgCl`, `PTC_WgLg` tracked green façade areas; `PTC_Wr` tracked the reference wall.

Weather Station (WS): recorded rainfall, wind, temperature, and humidity ∼6 m away, providing meteorological context.

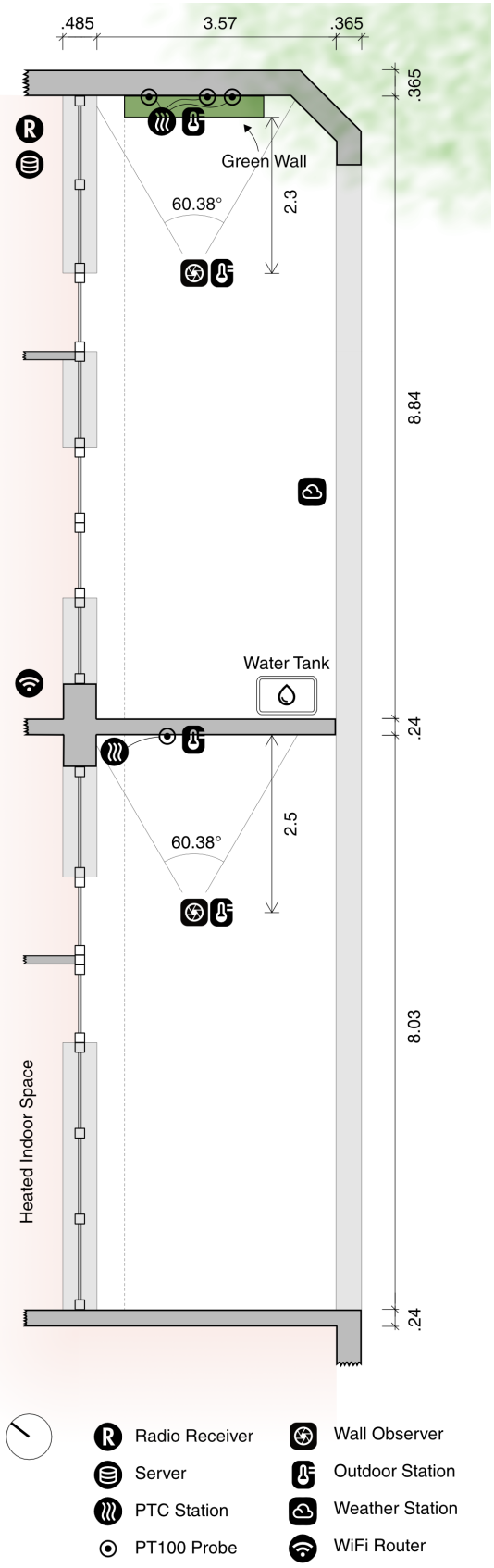


Figure 1: Floorplan of the terraces, dimensions in meters

2.3 Data Acquisition and Networking

A local Wi-Fi router collected sensor data from Tinkerforge ESP32 Bricks, with a laptop serving as the main server for Python-managed data storage. Figure 3 depicts the data flow from sensors to databases. Data collection occurred at ∼5-minute intervals for all parameters, utilizing sub-second in-memory buffering to mitigate hardware latency and packet loss. Storage intervals were structured by level-of-detail (LOD) requirements for analysis robustness and data yield optimization. We implemented three LOD scales: **LOD/cache** (sub-second) for in-memory buffering; **LOD/realtime** (5 s, ±1 s tolerance) for immediate visualization and system health monitoring, with 3-hour retention in a time series database; **LOD/day** (5 min, ±30 s tolerance) for daily analysis in persistent storage (mongoDB). Additional LOD scales for weekly, quarterly, or decadal storage were conceptualized but not implemented.

Figure 2: Left to right: Wall Observer Station, Outdoor Station, PTC Station, Weather Station

2.4 SMDP Format

Measurement data was sanitized into a *Small Measurement Data Packet (SMDP)* JSONL format. Each line contains an object with four keys: (1) `"i"`: station ID; (2) `"t"`: timestamp in ISO 8601 format; (3) `"p"`: parameter name; (4) `"v"`: measured value (unitless). This minimalistic approach optimizes file size and parsing efficiency for iterative analysis workflows, particularly when processing datasets multiple times with matplotlib and similar tools. While CSV offers a viable alternative, SMDP provides superior flexibility for varied parameter types, with contextual metadata maintained in separate files.

2.5 Data Management

Python scripts operated in isolated processes to prevent interference with Tinkerforge's internal callback loops. Data transfer from short-term time-series to long-term database collections required generating new IDs to prevent MongoDB insertion conflicts. System stability improved significantly through enhanced error handling implemented from data batch 3 onward (see Fig.4).

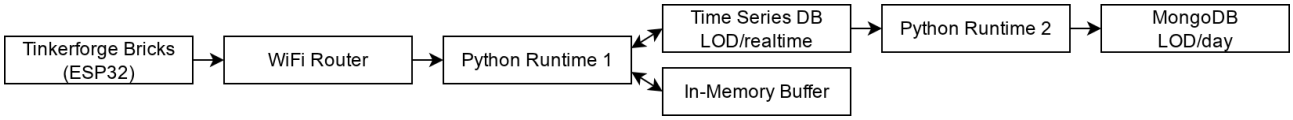


Figure 3: Conceptual diagram of data flow, in-memory buffering, and multi-stage database storage. Arrows represent data flow and approximate network paths.

3 Results

3.1 Data Batches and Yield

Over 127 days (16 May–20 Sep. 2024), four measurement batches (B1–B4) exceeded two-day duration with over 50% data yield, calculated relative to optimal 5-minute intervals across all sensors. Yield rates improved progressively through resolution of connectivity and database issues (Figure 4). This analysis focuses on batch B3 (29 July–14 Aug. 2024, 15.4 days), which achieved 82.8% average yield with complete sensor functionality until one wall observer disconnected near the batch conclusion.

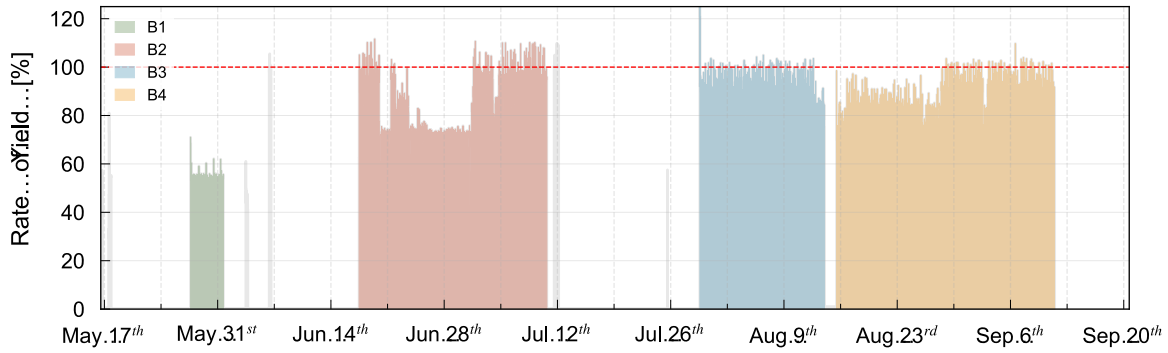


Figure 4: Data yield rates for measurement batches B1–B4. Progressive improvement in B3 and B4 reflects resolution of connectivity and database issues.

3.2 Thermal Comparison

Weather station data (Figure 5) shows ambient temperatures ranging from 15 °C to 38 °C during B3, while outdoor stations (Figure 6) recorded significantly higher values reaching 50 °C. This discrepancy stems from the weather station's radiation shield versus unshielded outdoor sensors—a critical design oversight affecting data interpretation.

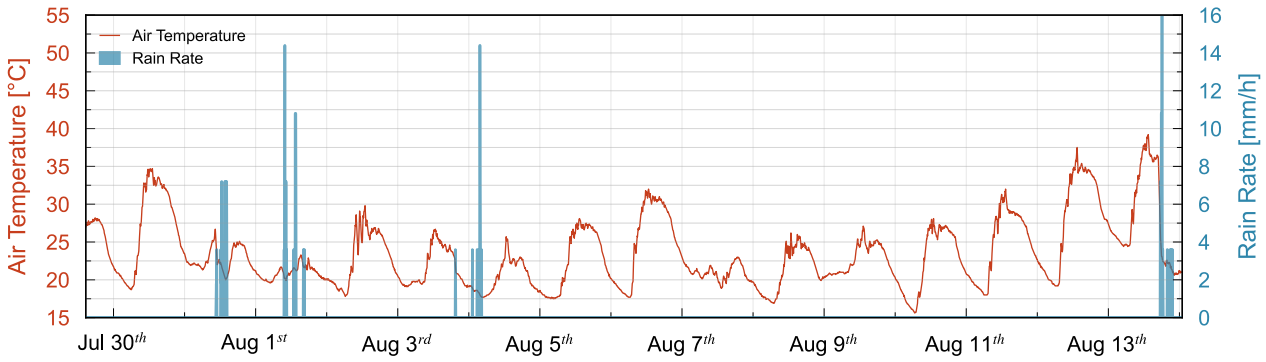


Figure 5: Rain and temperature measurements at the weather station.

Despite measurement limitations, air temperatures at the green wall consistently measured 2K cooler than the reference wall throughout B3. Wall observer temperatures remained nearly identical until August 13th, when reference wall readings shifted to match green wall temperatures—likely due to sensor ID reassignment during reconnection. This problematic design behavior, confirmed

in subsequent deployments, only affects the OS and WS stations in conjunction with Tinkerforge's Outdoor Weather Bricklet (not programmable by their API).

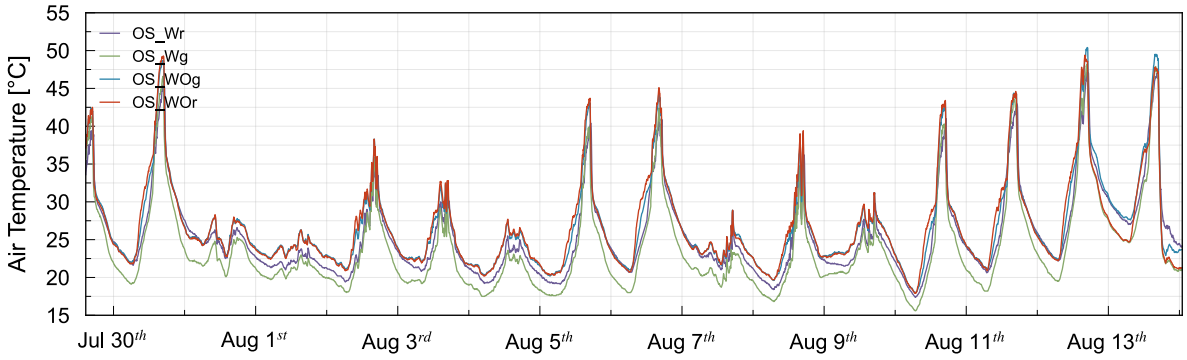


Figure 6: Comparison of air temperatures on both terraces: `OS_Wg` @ green wall, `OS_Wr` @ reference wall, `OS_WOg` @ green wall observer, `OS_WOr` @ reference wall observer.

Radiant temperature measurements (Figure 7) reveal radiation spikes on the reference wall observer attributable to geometric shading differences (Figure 8) rather than façade effects. Both half-sphere sensors measured similar baseline radiation when shading effects are excluded.

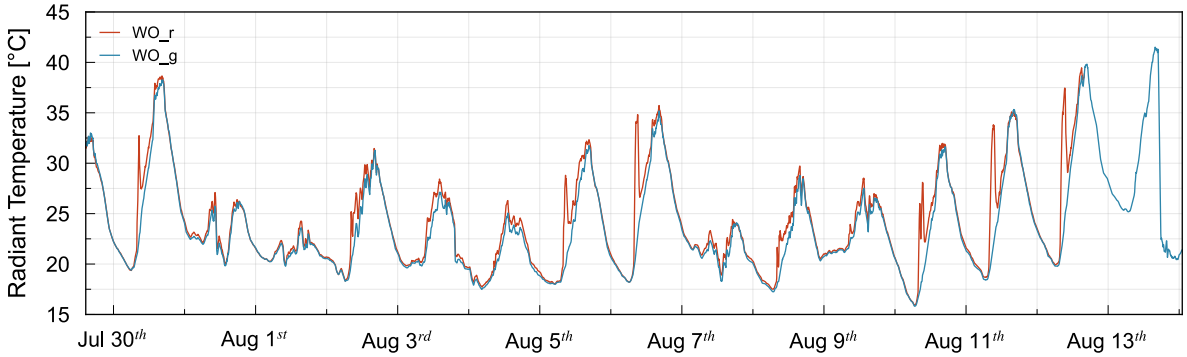


Figure 7: Radiant temperature (black-bulb sensor) from the `WO_g` (green wall observer) compared to the `WO_r` (reference wall observer).

Figure 8: Hourly simulated difference in shading conditions (including tree) from 05:00 to 21:00 (raytraced). First row: Vegetated wall, second row: Reference wall. As seen from the center of each terrace. This visualization was done after the experiment to provide context.

These findings demonstrate measurable green façade cooling effects despite experimental limitations, highlighting essential considerations for future sensor deployment protocols.

Figure 9: Placement of PT100 temperature probes behind the BGK, affixed to the wall.

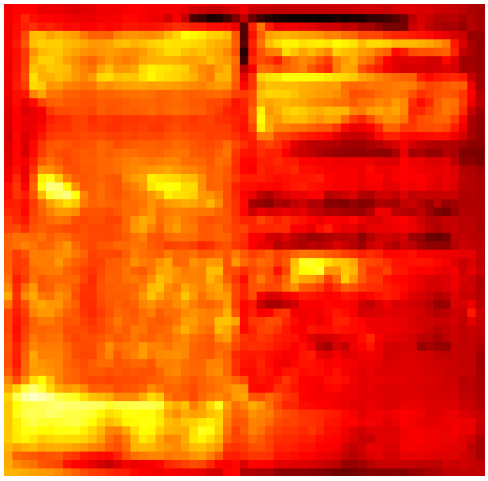


Figure 10: Representative thermal image: BGK / B3 / August 5th, 2024 at 14:27:30.

3.2.1 Thermal Imaging

A primary objective was characterizing thermal behavior differences between the green façade and reference wall through continuous thermal imaging. Figure 11 presents a thermal difference matrix (green - reference) containing 1,622,400 temperature measurements across B3. This continuous imaging approach provides significantly more detailed spatial and temporal insights than point-based black bulb measurements.

The thermal images reveal complex surface temperature patterns. The BGK's black fleece textile registers temperatures up to 10-15K higher than the reference wall during peak solar exposure, while vegetated portions demonstrate cooling effects up to 10K cooler. Cooling effects are most pronounced during afternoon and nighttime periods, with thermal equilibrium around 8-9 AM before the green façade becomes relatively cooler again by 11 AM.

This hourly analysis across multiple days demonstrates the value of continuous thermal imaging for capturing diurnal thermal cycles and identifying optimal periods for green façade cooling performance. The spatial resolution allows identification of specific surface components contributing to thermal differences, providing insights unattainable through traditional point sensor measurements.

4 Discussion and Conclusions

4.1 Methodological Insights and Recommendations

This deployment revealed key lessons for sensor reliability and data management. **Sensor standardization:** All sensors measuring identical parameters require consistent configurations—outdoor air temperature sensors need radiation shields, explaining the 50 °C vs. 38 °C discrepancy between unshielded outdoor sensors and the weather station. **Measurement approach:** Converting surface temperatures to incoming radiant flux (watts) would provide more robust thermal environment comparisons by eliminating emissivity and distance assumptions inherent in thermal camera measurements [5]. **System reliability:** Sub-second in-memory caching successfully compensated for callback latencies, while MongoDB proved unnecessarily complex—SQLite would suffice. Comprehensive exception

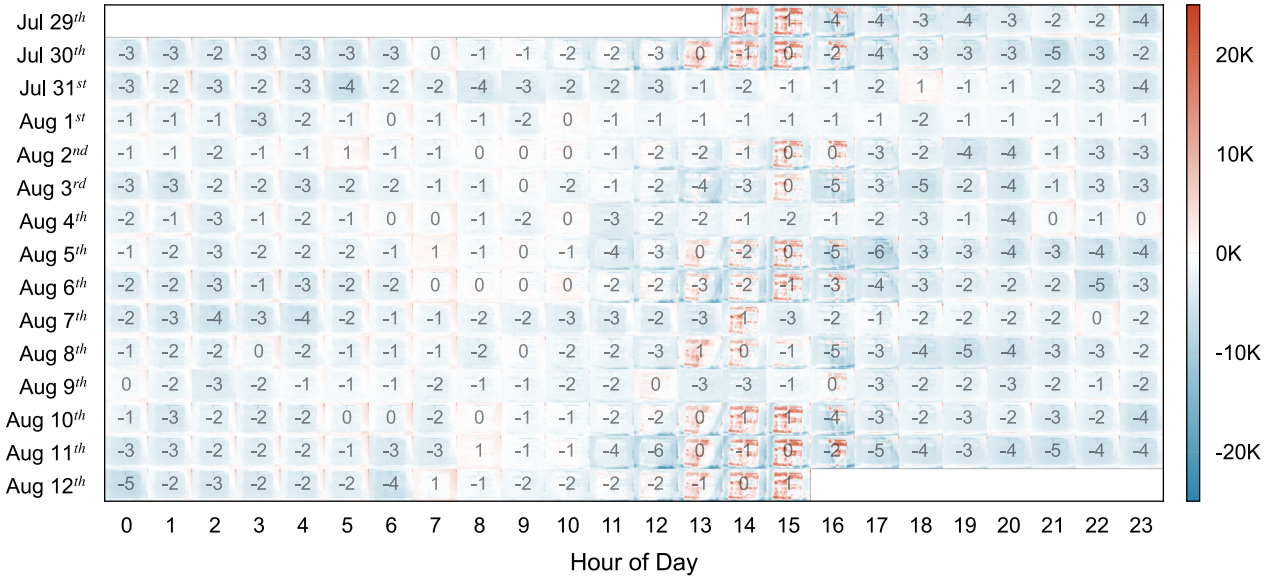


Figure 11: Thermal images showing temperature differences between green and reference surfaces. Each image displays average hourly temperature difference overlays, revealing diurnal thermal patterns and spatial temperature variations.

handling is crucial to prevent system crashes during extended deployments. **Site analysis:** Thorough environmental assessment must account for wind patterns (behind green facade wall), thermal mass (differing wall thickness), seasonal vegetation changes (deciduous tree), and infrastructure interference (water tank) before sensor placement. Cross-comparing both terraces by relocating the BGK could address these site-specific variations.

4.2 Conclusions

We successfully deployed a low-cost sensor network demonstrating measurable green façade cooling effects, particularly during nighttime periods. This multi-parameter approach provides a template for monitoring nature-based urban solutions. Future research should: (1) perform comprehensive site analysis, (2) standardize sensor configurations, (3) measure radiant flux rather than surface temperatures, (4) implement spherical measurements for complete thermal radiation assessment, and (5) expand to year-round monitoring. Despite design flaws and unexpected challenges arising continuously, we believe these experiences provide valuable insights for the research community. Our most significant contribution may be demonstrating continuous outdoor thermal imaging over extended periods—a capability not previously documented in green façade research.

5 Acknowledgments

We thank the Chair of Landscape Architecture—specifically Axel Timpe—for initiating this research project and the research design, entrusting us with developing the monitoring setup, and providing support throughout the design and experimental phases.

Yuxia Luo (Chair of Landscape Architecture) made the most substantial individual contribution by co-designing the measurement stations, particularly the PTC sensor nodes, and conducting regular on-site inspections and troubleshooting throughout the deployment period.

We are grateful to Tinkerforge for their exceptional technical support, including code examples and expert advice that facilitated effective use of their sensor API.

We also acknowledge students from the Chair of Gebäudetechnologie (GBT) who conducted preliminary analyses as part of the research seminar "Fassadenbegrünung auf dem Prüfstand" (Green Façades Under Scrutiny) during the project's initial phase. Their insights informed our subsequent data analysis approach, and we thank Felix Schürfeld for providing the photographs featured in Figure 2.

Equipment and student assistance for this research have been funded as part of the BauGrünKit project at the Chair of Landscape Architecture by RWTH Aachen University's Sustainability Fund.

References

[1] I. Blanco and F. Convertino, "Thermal performance of green façades: Research trends analysis using a science mapping approach", *Sustainability*, vol. 15, no. 13, p. 9981, 2023. DOI: 10.3390/su15139981 [Online]. Available: https://www.mdpi.com/2071-1050/15/13/9981

[2] M. Martin, A. Chong, F. Biljecki, and C. Miller, "Infrared thermography in the built environment: A multi-scale review", *Renewable and Sustainable Energy Reviews*, vol. 165, p. 112 540, 2022. DOI: 10.1016/j.rser.2022.112540 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032122004397

[3] A. Timpe et al., "BauGrünKit Green Rack – Entwicklung und Erprobung eines Selbstbaukits für die koproduktive Begrünung von Hochschulgebäuden", RWTH Aachen, Institut und Lehrstuhl für Landschaftsarchitektur, Aachen, Tech. Rep., 2023, p. 14. DOI: 10.18154/RWTH-2023-09121 [Online]. Available: https://publications.rwth-aachen.de/record/969337/files/969337.pdf

[4] Tinkerforge GmbH, *Tinkerforge documentation*, Tinkerforge GmbH, 2024. Accessed: Aug. 2, 2025. [Online]. Available: https://www.tinkerforge.com/en/doc/

[5] J. Loveday, G. Loveday, J. Byrne, B.-l. Ong, and P. Newman, "Quantifying radiation from thermal imaging of residential landscape elements", *Renewable Energy and Environmental Sustainability*, vol. 2, p. 17, 2017. DOI: 10.1051/rees/2017041 [Online]. Available: https://www.rees-journal.org/articles/rees/full_html/2017/01/rees170041s/rees170041s.html

Enhancing Trust in Machine Learning-Based Architectural Design Space Exploration Through Uncertainty Visualization, a Pilot Study

Max Zorn[1] and Thomas Wortmann[1]
[1]Department for Computing in Architecture, Institute for Computational Design and Construction, University of Stuttgart, Stuttgart, Germany
E-mail(s): max-benjamin.zorn@icd.uni-stuttgart.de, thomas.wortmann@icd.uni-stuttgart.de

Abstract: Design Space Exploration (DSE) tools enable designers to assess numerous design variants without the extensive computational demand of traditional simulations by utilizing Machine Learning (ML) methods. These advanced ML-based tools empower designers by providing real-time feedback and allowing interactive evaluation of promising design candidates. However, the inherent uncertainties and potential errors associated with such models often lower the designers' trust, hindering a broader application in practice. This paper presents the application of Gaussian Processes to model and visualize the uncertainties inherent in ML-based DSE tools, evaluating whether visualizing model uncertainties, specifically through confidence bounds, positively affects designer trust. We conducted a pilot within-subject user study with architectural designers, investigating how their understanding of these uncertainties influences trust and decision-making. We compare three visualization modes with different levels of embedded information and assess the designers' perceived trust, decision-making confidence, and overall usefulness of each visualization in supporting exploratory tasks. We find that designers prefer visualizations that include uncertainty information, reporting higher trust, confidence, and decision-making effectiveness compared to modes lacking such cues.

Keywords: Design Space Exploration(DSE), Machine Learning, Data Visualization, Explainable AI, Trust

DOI: 10.18154/RWTH-CONV-254899. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Architectural DSE has increasingly benefited from ML methods, enabling designers to quantitatively assess numerous design variants ad hoc without the extensive computational demand of traditional simulation methods. Such advanced ML-based tools empower designers by providing real-time feedback through pre-trained prediction models and data visualization, allowing interactive evaluation of promising design candidates. A growing number of DSE tools are available as plugins for widely-used planning software such as *Rhino*®, but optimization and ML-based workflows still have not found

widespread adoption in planning practice [1] . The black-box nature of prediction models and their inherent errors lower the user's trust in the estimations [2]. The initial sampling methods can further significantly impact the accuracy of prediction models. For example, models trained on optimization data might be accurate locally in areas of expected high performance but lack global accuracy. In contrast, models trained on uniformly distributed samples show higher global accuracy, but may fail to depict areas of high performance [3]. Unlike algorithmic sampling methods like optimization, DSE tools give agency to the human designers to decide which design candidates to analyze in more detail. Crucially, DSE tools present an estimated performance and simultaneously allow the designers to verify the performance by simulating a chosen candidate, if they do not want to rely solely on the estimation. However, two problems arise from this feedback loop. Firstly, the users are unaware of the model's inherent accuracy, prompting the question of whether or not it's worth simulating a chosen point, specifically if simulations are expensive. Secondly, repeated observations of high model errors significantly lower the user's trust [2], ultimately leading to users disregarding the predictions or stopping the use of the tool.

In this paper, we present an initial study of how adding information about the model's uncertainties to the visualization can circumvent these shortcomings. We conducted a small-scale user study and found that visualizing model uncertainty enhances user trust and affects their exploration strategy.

1.1 Background

While many DSE tools from other application domains exist, we only present a selection of DSE tools that are developed for the AEC community and some studies from the human-computer interaction domain that focus on trust in ML, and visualizing and interacting with uncertainties.

Wortmann et al. [4] introduce *Performance Explorer*, an interactive ML-based DSE tool that visualizes the fitness landscape as a 2D heatmap using star-coordinates, using surrogate models from RBFOpt to estimate performance. Integrated into the *Opossum* plugin for *Grasshopper*, it allows designers to explore designs interactively. Zorn [3] and Abdelaal et al. [5] extend this tool. The former analyzed the influence of different sampling and ML methods on the accuracy of such maps, the latter introduced a discrete glyph-based visualization of the maps, visually encoding additional information, such as the parameter configuration of each point.

While these approaches use regression models for forward DSE, Conditional Variational Autoencoder (CVAE)s enable inverse workflows. Danhaive et al. [6] and Balmer et al. [7] employ CVAEs to learn performance-conditioned design spaces. Danhaive et al. [6] use performance-guided sampling to better capture the well-performing areas of the design space. Balmer et al. [7] integrate explainable ML techniques like sensitivity analysis and decision trees to support interactive latent space exploration.

Berger et al. [8] focus on enabling uncertainty-aware exploration of continuous parameter spaces via multivariate prediction, combining real-time surrogate modeling with visual encodings of predictive uncertainty to guide interactive sensitivity analysis. Broader surveys by Jena et al. [9] and Weiskopf [10] synthesize visualization strategies for uncertainty, highlighting ongoing challenges in representation, propagation, and perception of uncertainty across the visualization pipeline, especially for decision-critical domains.

2 Methods

To understand the influence of visualizing model uncertainties on the user's trust, we build a simple interactive, ML-based visualization prototype (Figure 1) that offers three visualization modes (Figures 2-4), with different levels of displayed information. We use this prototype to conduct a small-scale within-subject user study. Participants are tasked to explore the abstracted design space and identify well-performing points for each visualization mode. Every visualization mode has a different ground truth function, which calculates the performance (i.e., the y-value).

The following subsections will explain the tool and study design in more detail.

2.1 Gaussian Process Regression

As the name implies, Gaussian Process Regression (GPR) is a statistical modeling approach that provides point estimates and a full probability distribution over possible outcomes. This allows for quantifying uncertainty in the model's predictions, typically represented as a confidence interval. GPR achieves this by assuming a prior distribution over functions and updating this prior using observed data through a process akin to Bayesian interpolation. In regions where the model has limited data, the confidence interval tends to widen, indicating greater uncertainty, while it narrows in areas where observations are dense. Therefore, the confidence interval serves as an indicator of the model's inherent uncertainty in its predictions.

We implement a custom GPR in *Python* using *SciPy* [11] and *NumPy* [12], and define a custom composite kernel that combines a scaled Radial Basis Function (RBF) with a linear kernel and a constant bias:

$$k(x, y) = \tau \cdot \exp\left(-\frac{(x-y)^2}{\ell}\right) + 0.25 \cdot xy^\top + 0.15 \tag{1}$$

where τ is the amplitude parameter of the RBF and ℓ is its length-scale parameter.

During the interaction, the GPR is fitted to the observations, starting with only one point at $x = 0$, as seen in Figure 1. Also, notice how the confidence interval grows with distance from this point and collapses near the observation.

2.2 DSE Tool

The interactive, browser-based DSE prototype was developed in *Python* using the *Streamlit* library for the interface and *Plotly* [13] for interactive visualizations. The prototype layout (Figure 1) consists of a left-hand sidebar and a main display window. The sidebar includes navigation between the three visualization modes (Area A) and study-relevant controls (Area B), where users can toggle the ground truth function and access a short description of the visualization and displayed data. The main display window contains the visualization itself (Area C), along with the slider and action buttons for adding observations or proceeding to the next stage of the study (Area D).

Figures 2–4 illustrate the three visualization modes implemented in the prototype. In the *Manual Mode*, only the participant's added observations are displayed, with no additional information about the remainder of the design space or the likely performance of unobserved points. The *Predictive Mode* displays observations in the same way but adds a regression line representing the mean of the

Figure 1: Interface of the DSE prototype in *Uncertainty Mode*. (A) Navigation, (B) controls and description, (C) visualization with observations, regression line, and confidence interval, (D) slider and task controls.

predictive distribution from the GPR. The *Uncertainty Mode*mode builds on the *Predictive Mode* by also showing the GPR confidence bounds. In all modes, users may optionally display the ground truth function, shown as a blue line in Figure 5. This functionality was disabled during the study task (see Subsection 2.3).

2.3 Study Design

We recruited twelve participants: five young professionals and seven students, all under 35 years of age. Seven identified as male, three as female, one as diverse, and one preferred not to disclose gender. Their professional backgrounds included architecture (n=6), civil engineering (n=3), data science (n=2), and electrical engineering/UX design (n=1). Five participants had more than five years of professional experience. All reported some prior interaction with data visualization. The study was conducted in English at the participants' work desks.

Each participant interacted with the three visualization modes described in Subsection 2.2, presented in randomized order. In each mode, the task was to identify points with low y-values, where y represents the maximum deflection of a slab. This setup reflects DSE methods, but without visual cues indicating what a high- or low-performing design candidate might look like.

At the start of a task, the visualization contained only one pre-computed solution at $x = 0$ (see Area C in Figure 1). Participants could add up to five observations by selecting an x-value from the range $x \in [-5, 5]$ using a slider and then pressing the *Add Observation* button. Alternatively, they could end the task earlier by pressing the *Done* button. Once either condition was met, the questionnaire for that mode appeared below the visualization, and the *Show/Hide Ground Truth* buttons became available.

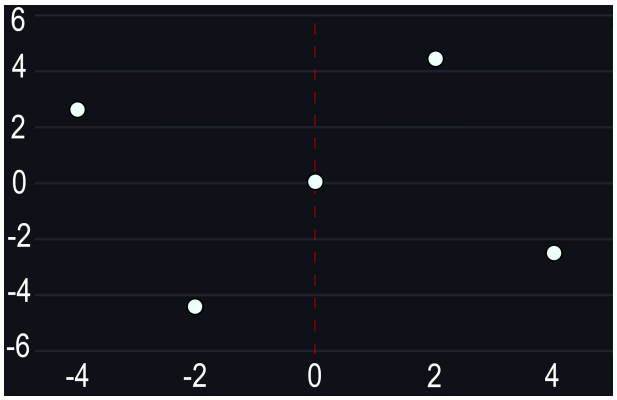


Figure 2: *Manual Mode*: Observations are visualized as points.

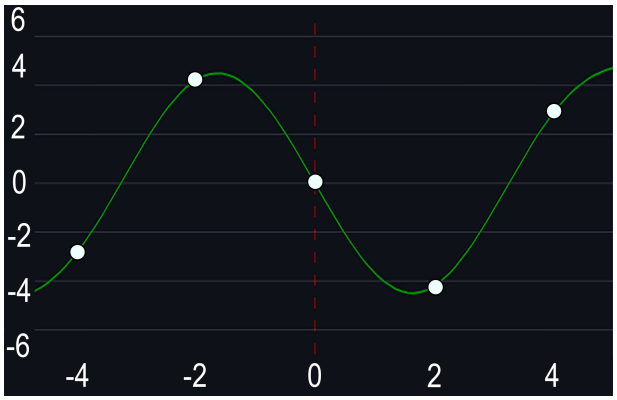


Figure 3: *Predictive Mode*: Observations (points) and regression line (green).

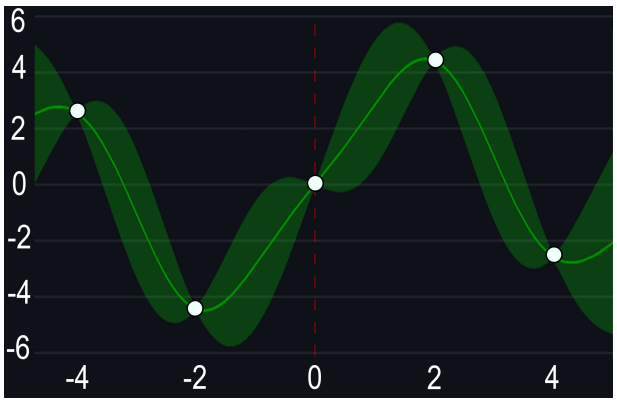


Figure 4: *Uncertainty Mode*: Observations (points), regression line (green), and confidence interval (green area).

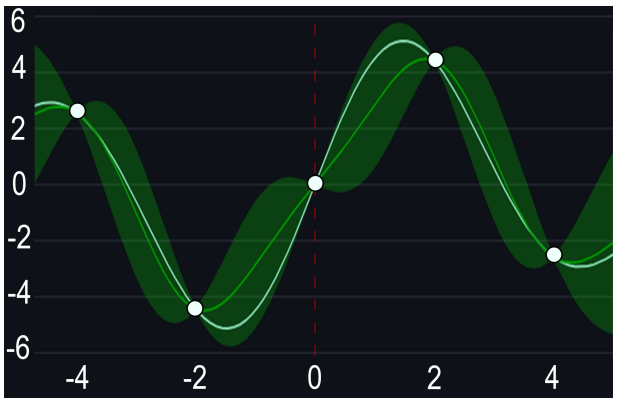


Figure 5: *Ground Truth*: Additional visualization of true function (blue line).

After completing the questionnaire for a given mode, participants moved on to the next visualization. This process was repeated until all three modes had been used. Following these interactions, participants completed a final set of evaluative questions and could provide written feedback through a text input field, allowing them to elaborate on their experiences and impressions of the visualizations.

3 Results

We present results from post-visualization and final assessment questionnaires, both using 5-point Likert scales (1 = Strongly Disagree, 5 = Strongly Agree).

Figure 6 shows responses to six items across the three visualization modes. Boxplots illustrate response distributions for direct comparison. Across all items, the *Uncertainty Mode* received the most favorable ratings, particularly for perceived support, effectiveness, and preference for future use. *Manual Mode* consistently had the lowest medians and greater variability.

For *"I felt confident selecting points based on this visualization"*, *Uncertainty Mode* scored highest ($\mu = 4$), followed by *Predictive Mode* ($\mu = 3$) and *Manual Mode* ($\mu = 2$). The same trend appeared for model trust (*Uncertainty Mode* $\mu = 4$, *Predictive Mode* $\mu = 3$, *Manual Mode* $\mu = 2$), with *Manual Mode* showing wider response spread. Participants also rated *Uncertainty Mode* highest for identifying high-

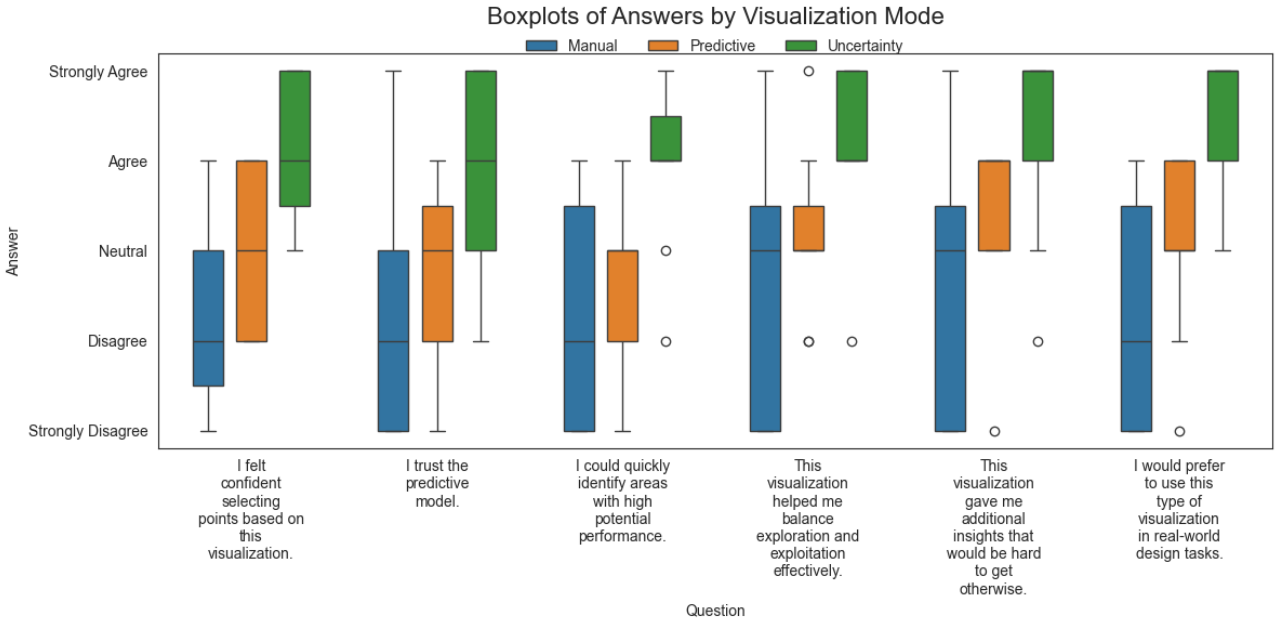


Figure 6: Distribution of responses to post-visualization questionnaires.

potential performance areas ($\mu = 4$), while *Predictive Mode* and *Manual Mode* tended toward neutral or disagree. *Manual Mode* responses often had wide interquartile ranges, suggesting inconsistent experiences (e.g., lucky guesses).

For balancing exploration/exploitation and generating novel insights, *Uncertainty Mode* again outperformed, with medians of 4–5 versus <4 for the other modes. When asked about real-world use, participants clearly preferred *Uncertainty Mode* ($\mu = 4$). Overall, uncertainty-inclusive visualizations were perceived as more trustworthy, insightful, and actionable, better supporting decision-making than the *Manual Mode* or *Predictive Mode*.

After all three modes, participants answered final evaluative questions (Figure 7). Most preferred more detailed visualizations, with no negative responses to *"I would prefer more detailed visualizations in similar tasks"*. Trust increased with added information—most agreed or strongly agreed with *"I felt increasing trust as more model information was visualized"*, and none disagreed.

Uncertainty visualization also improved decision-making: for *"Seeing uncertainty (if present) improved my decision-making"*, all responses were positive or neutral, with agreement being most common. Most participants also reported adapting their exploration strategy based on the visualized information. When asked which visualization was most helpful for selecting high-performing points, responses were largely positive, though *Manual Mode* again showed mixed reactions, consistent with its broader response ranges above.

4 Discussion and Outlook

The study presented in this paper shows that including uncertainty in data visualizations significantly improves user trust, decision-making, and satisfaction. Across tasks, participants consistently favored the *Uncertainty Mode* over the *Manual Mode* and *Predictive Mode*. Final evaluations reinforced these

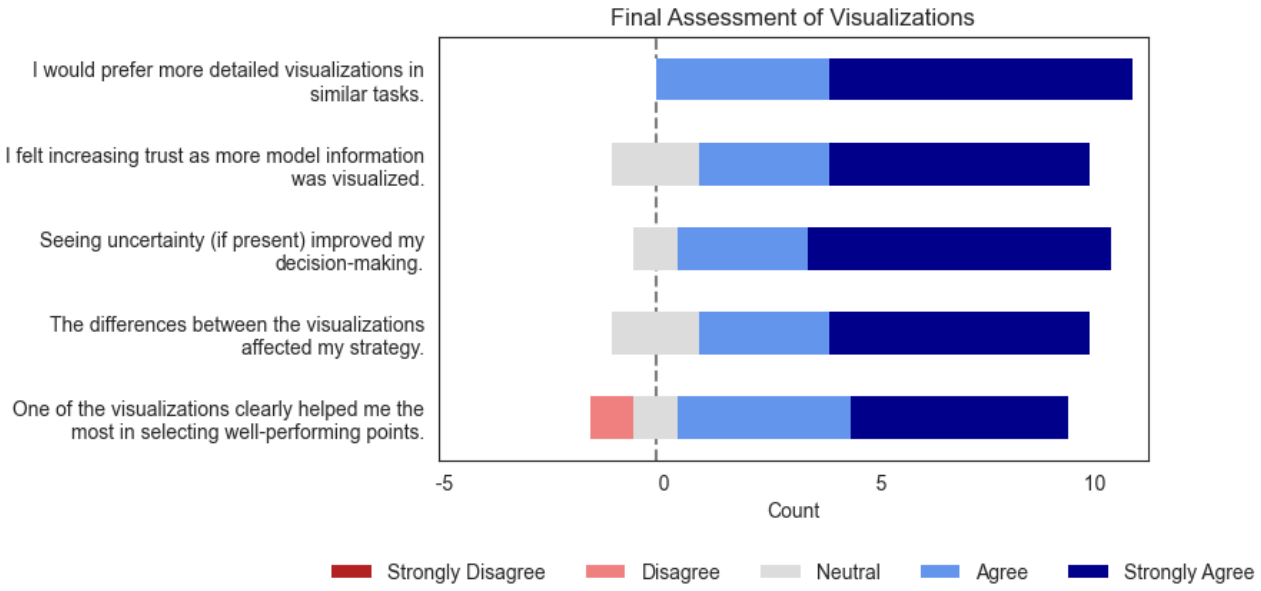


Figure 7: Likert scale of questions answered after interacting with all visualization modes.

trends: most participants said that uncertainty information improved their decisions and increased their trust. Many also expressed a desire for more detailed visualizations in future tasks. Participants also adapted their exploration strategies based on the visualization mode. However, our study has a couple of limitations. As a pilot study by design, the number of participants was relatively low, and the tested system was simplified. Unlike real-world DSE tasks, which are high-dimensional and highly non-linear, the designers explored and interacted with a low-dimensional, linear system. Furthermore, only one visual encoding of uncertainty was tested.

Future work should therefore focus on testing a variations of visual encodings of uncertainty and study how this interaction affects user understanding and task performance on real-world (architectural) design tasks.

Acknowledgments

This research was partially funded by Horizon Europe and the European Innovation Council and SMEs Executive Agency (EISMEA) via the EIC Pathfinder Project “UniversalTimberSlab” with GA Project Number 101161103. This research was partially supported by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) under Germany’s Excellence Strategy - EXC 2120/1 - 390831618.

References

[1] T. Wortmann, J. Cichocka, and C. Waibel, “Simulation-based Optimization in Architecture and Building Engineering—Results from an International User Survey in Practice and Research”, *Energy and Buildings*, vol. 259, p. 111 863, 2022.

[2] M. Yin, J. Wortman Vaughan, and H. Wallach, “Understanding the Effect of Accuracy on Trust in Machine Learning Models”, in *Proceedings of the 2019 CHI Conference on Human Factors in*

Computing Systems, Glasgow Scotland Uk: ACM, May 2019, pp. 1–12. DOI: 10.1145/3290605.3300509

[3] M. Zorn, “A novel software framework for architectural design space exploration”, in *Proceedings of the 34th Forum Bauinformatik 2023*, Bochum: Ruhr-Universität Bochum, 2023, 8 pages. DOI: 10.13154/294-10106

[4] T. Wortmann, T. Schroepfer, and Z. Akbar, “Surveying Fitness Landscapes with Performance Explorer - Supporting the Design of a Better Tomorrow with Interactive Visualizations”, in *Proceedings of the 36th eCAADe Conference*, A. Kępczyńska-Walczak and S. Białkowski, editors, vol. 1, Lodz, PL: eCAADe and Faculty of Civil Engineering, Architecture and Environmental Engineering, Lodz University of Technology, 2018, pp. 621–630.

[5] M. Abdelaal et al., “Visual analysis of fitness landscapes in architectural design optimization”, *The Visual Computer*, vol. 40, Jun. 2024. DOI: 10.1007/s00371-024-03491-3

[6] R. Danhaive and C. T. Mueller, “Design subspace learning: Structural design space exploration using performance-conditioned generative modeling”, *Automation in Construction*, vol. 127, p. 103 664, Jul. 2021. DOI: 10.1016/j.autcon.2021.103664

[7] V. Balmer et al., “Design Space Exploration and Explanation via Conditional Variational Autoencoders in Meta-Model-Based Conceptual Design of Pedestrian Bridges”, *Automation in Construction*, vol. 163, p. 105 411, Jul. 2024. DOI: 10.1016/j.autcon.2024.105411

[8] W. Berger, H. Piringer, P. Filzmoser, and E. Gröller, “Uncertainty-Aware Exploration of Continuous Parameter Spaces Using Multivariate Prediction”, *Computer Graphics Forum*, vol. 30, no. 3, pp. 911–920, Jun. 2011. DOI: 10.1111/j.1467-8659.2011.01940.x

[9] A. Jena, U. Engelke, T. Dwyer, V. Raiamanickam, and C. Paris, “Uncertainty Visualisation: An Interactive Visual Survey”, in *2020 IEEE Pacific Visualization Symposium (PacificVis)*, Tianjin, China: IEEE, Jun. 2020, pp. 201–205. DOI: 10.1109/PacificVis48177.2020.1014

[10] D. Weiskopf, “Uncertainty Visualization: Concepts, Methods, and Applications in Biological Data Visualization”, *Frontiers in Bioinformatics*, vol. 2, p. 793 819, Feb. 2022. DOI: 10.3389/fbinf.2022.793819

[11] P. Virtanen et al., “SciPy 1.0: Fundamental algorithms for scientific computing in Python”, *Nature Methods*, vol. 17, no. 3, pp. 261–272, Mar. 2020. DOI: 10.1038/s41592-019-0686-2

[12] C. R. Harris et al., “Array programming with NumPy”, *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020. DOI: 10.1038/s41586-020-2649-2

[13] P. T. Inc., *Collaborative data science*, https://plot.ly, Montreal, QC, 2015.

Comprehensive Georeferencing Support for all Official Versions of the IFC Standard

Štefan Jaud[1], Paul-Christian Max[2], Torben Pullmann[3] and Andreas Geiger[4]
[1]Jaud IT GmbH, Moorenweis, Germany
[2]AFRY Deutschland GmbH, Leipzig, Germany
[3]albert.ing GmbH, Frankfurt am Main, Germany
[4]Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
E-mail(s): stefan@jaud-it.com, paul-christian.max@afry.com, t.pullmann@albert-ing.com, andreas.geiger@kit.edu

Abstract: Industry Foundation Classes (IFC) is a standardized vendor-neutral data schema that enables exchange of digital data between different stakeholders in the architecture, engineering and construction (AEC) industry. Last year, a new version of the standard was published – so called IFC4.3 – which adds new entities and concepts to support various georeferencing scenarios (ISO 16739:1-2024). While this enables crisp data exchanges (as soon as the standard is adopted by the software vendors), the question arises about the previous official versions of IFC: can these be updated in a non-intrusive manner to support the scenarios as well? This contribution proposes an extension in the form of specialized property sets attached to the top node of the IFC data structure. The resulting "add on" Model View Definition (MVD) have been prototypically implemented in two independent software solutions with positive results. Additionally, example files were produced to showcase the different possibilities of georeferencing, which shall enable streamlined implementation for other software products.

Keywords: Building Information Modeling (BIM), georeferencing, Industry Foundation Classes (IFC), Model View Definition (MVD)

DOI: 10.18154/RWTH-CONV-254900. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

1.1 Problem Statement

The new IFC data schema IFC 4.3.2.0 was published recently [1]. With it, an extension to the georeferencing vocabulary was introduced, which allows for a more thorough support of the various scenarios occurring in the industry [2], [3]. Thus, the standard enables semantically crisp IFC exchanges by providing the possibility to exchange detailed georeferencing metadata for a wider range of building information modelling (BIM) projects in the AEC industry [4].

However, there is a timing problem: the new standard has to be adequately widely adopted by the many software vendors supporting BIM processes before such exchanges can take place [5]. Thus, there is a need to explore the possibilities for adoption of the new options using the existing IFC interfaces implemented in the established software products. As such, a question arises about the previous official versions of the IFC standard: *can these be updated in a non-intrusive manner to support the new scenarios as well?* In other words, can a workaround be defined such that there would be little-to-no programming effort required to employ existing implementations?

Let it be noted here, that we are interested in *full* georeferencing. This includes data about the underlying coordinate reference system (CRS) together with the coordinate operation (CO), i.e. model's insertion point as well as its rotation within a chosen CRS. This corresponds to level of georeferencing (LoGeoRef) 50 as defined in [6].

Related to our work, [7] has already proposed a workaround for the possibilities introduced with IFC4 and how to encode these in IFC2x3 as explained in detail in Section 2.3. There has been no similar studies or proposals that address novelties of IFC4.3 to the best of authors' knowledge, which represents the research gap explored in this paper.

1.2 Methodology and the Structure of the Paper

The research was conducted following the Design Science Research (DSR) method as established by [8]. The DSR method focuses on creation and evaluation of Information Technology (IT) artefacts which solve an identified problem. The method consists of the following steps [8]:

1. problem identification and motivation,
2. definition of objectives for a solution,
3. design and development of artefacts solving the problem,
4. demonstration of suitability of developed artefacts to solve the problem,
5. evaluation of the solution by comparing the objectives and obtained results, and
6. communication of the obtained results.

The paper is structured following the DSR method presented above. This section describes our motivation, states the problem and the objective for the solution. Next section describes the state of IFC data schema and workarounds regarding georeferencing support for all three official versions: IFC2x3, IFC4 and IFC4.3. Section 3 describes the proposed solution together with exemplary serializations. The suitability to solve the problem is demonstrated with two independent prototypical implementations in Section 4. We conclude with an evaluation of the results in Section 5.

2 Georeferencing in IFC Data Schema

The IFC data schema has been in development for the past 30 years. Since then, three versions were published by International Organization for Standardization (ISO) as a standard: IFC 2.3.0.1 in 2006 (i.e. IFC2x3), IFC 4.0.2.1 in 2013 with an update in 2018 (i.e. IFC4), and lastly IFC 4.3.2.0 in 2024 (i.e. IFC4.3) [1]. While the former does not offer any support for LoGeoRef50 as seen in Figure 1, the latter two added support in turn for the most occurring scenarios in vertical and horizontal BIM modelling practices. Sections 2.1 and 2.2 explain the individual concepts as introduced by IFC4 and

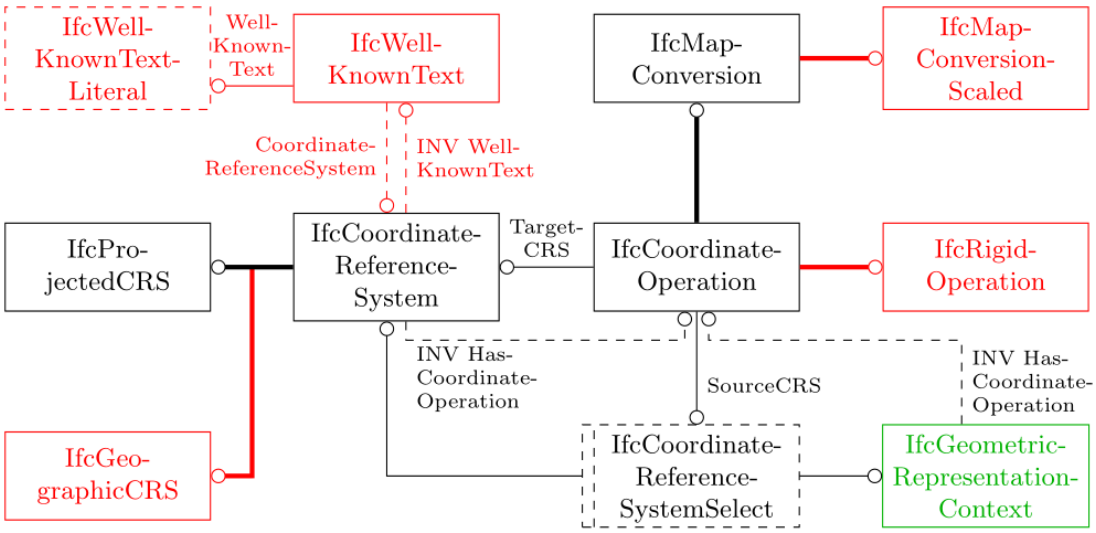


Figure 1: EXPRESS-G diagram of georeferencing entities in all IFC versions. Green is the state of IFC 2.3.0.1, black are additions within IFC 4.0.2.1; red are additions in IFC 4.3.2.0 [updated from 3].

IFC4.3, respectively, for various types of CRS and CO [9]. The existing proposal for a workaround in IFC2x3 in Section 2.3 is presented here as well.

2.1 Initial Support in IFC 4.0.2.1

Initial support for georeferencing metadata includes the possibility to locate and orientate the local engineering context of a BIM model within a well-established CRS. The entities are presented in black on Figure 1 and encapsulated in the *Project Global Positioning* concept template in the specification [1]. This template specifies a relation of a IFCGEOMETRICREPRESENTATIONCONTEXT (with all IFCSHAPEREPRESENTATION occurrences referencing it) to map coordinates of a given CRS.

A CRS is modelled in IFC with the abstract entity IFCCOORDINATEREFERENCESYSTEM. Its first and mandatory attribute NAME encodes the CRS's identifier from the well-established European Petroleum Survey Group (EPSG) registry [2], [10]. The DESCRIPTION attribute optionally gives a human readable description of the CRS. Two additional EPSG codes can be provided for the underlying geodetic and vertical datums in GEODETICDATUM and VERTICALDATUM attributes, respectively. A projected CRS is modelled with IFCPROJECTEDCRS and inherits from IFCCOORDINATEREFERENCESYSTEM. It encodes three properties of the used map projection: its name, its zone and the unit in attributes MAPPROJECTION, MAPZONE and MAPUNIT, respectively [1].

IFCMAPCONVERSION conveys that the IFCGEOMETRICREPRESENTATIONCONTEXT is a *topocentric* CRS. The first two parameters are inherited from IFCCOORDINATEOPERATION which connect a *source* CRS to a *target* CRS, e.g. the three-dimensional (3D) geometric context of BIM geometries with a projected CRS. The next three attributes (i.e. EASTINGS, NORTHINGS and ORTHOGONALHEIGHT) specify the coordinates of the source CRS's point of origin (PoO) in the target CRS. Following, XAXISABSCISSA and XAXISORDINATE define the orientation of the source's first coordinate axis within the target CRS. The source's and target's third coordinate axes coincide per definition. If these two attributes are omitted, the pairs of the first and second coordinate axes coincide as well. Last, the optional attribute SCALE allows for scaling between the used unit of measurement (UoM) in source

and target CRS, e.g. if BIM geometries are in feet and the underlying projected CRS is in meters. If omitted, the units are the same and the scale defaults to 1.

In a nutshell, the *Project Global Positioning* concept template defines the geometric context to be a *topocentric* CRS with the defined insertion point and horizontal rotation within the specified projected CRS. This is the most common modelling approach in the vertical BIM domains, where the BIM geometries in such context have their true dimensions in the model.

2.2 Comprehensive Extension in IFC 4.3.2.0

To support horizontal BIM modelling practices, a new entity IFCRIGIDOPERATION inheriting from IFCCOORDINATEOPERATION was introduced. In the case that TARGETCRS is of type IFCPROJECTEDCRS, this enables a mere translation of coordinates into the projected CRS. The attributes FIRSTCOORDINATE, SECONDCOORDINATE, and HEIGHT convey the offset values of BIM context's PoO in the projected CRS. As such, the geometric context of the BIM model shall be interpreted distorted equally to the underlying CRS.

To support constant scaling of a BIM model, a new entity IFCMAPCONVERSIONSCALED inheriting from IFCMAPCONVERSION was introduced to the IFC schema. It adds three additional scaling factors FACTORX, FACTORY, and FACTORZ, which enable the constant scaling of BIM dimensions along x, y, and z axes of the geometric context. This options represents a trade-off between irregularly distorted projected CRS and the size limit of topocentric CRS with its mismatch to any geospatial data. It is often used in infrastructure projects, e.g. the railway is planned in projected CRS, while for small structures along the way a constant scaling factor can be safely assumed.

Moreover, as put forward by [2], the use of a 3D geographic CRS for placement of topocentric PoO is enabled with the introduction of the new entity IFCGEOGRAPHICCRS. Its attributes GEODETICDATUM, PRIMEMERIDIAN, and UNIT establish the CRS following [9]. To connect the BIM geometries to the geodetic CRS, the IFCRIGIDOPERATION is allowed to support plane angle offsets in its attributes FIRSTCOORDINATE and SECONDCOORDINATE. Thus, IFCRIGIDOPERATION allows for both length and angle measures for IFCPROJECTEDCRS and IFCGEOGRAPHICCRS, respectively. Note that the HEIGHT attribute means ellipsoid height and not gravimetric height in this case.

Last but not least, IFC4.3 supports well-known text (WKT) strings to codify the CRS's parameters. This avoids the problem where an EPSG code is not available for the CRS used in BIM project. The WKT string is modelled with the IFCWELLKNOWNTEXTLITERAL type and connected to IFCCOORDINATEREFERENCESYSTEM with the IFCWELLKNOWNTEXT entity as seen on Figure 1. These entities allow to model the most common scenarios of georeferencing occurring in the AEC industry [11].

Table 1: An overview of all scenarios from [3] as fully supported in IFC4.3 with the combinations of IFCCOORDINATEREFERENCESYSTEM and IFCCOORDINATEOPERATION.

Scenario [3]	CRS	CO
Topocentric on geographic	IFCGEOGRAPHICCRS	IFCRIGIDOPERATION
Topocentric on projected	IFCPROJECTEDCRS	IFCMAPCONVERSION
Topocentric on averaged	IFCPROJECTEDCRS	IFCMAPCONVERSIONSCALED
Projected	IFCPROJECTEDCRS	IFCRIGIDOPERATION

2.3 Workaround for IFC 2.3.0.1

Realising the usefulness of the georeferencing additions in IFC4 when compared to IFC2x3, the international community proposed a workaround in the form of specialized property sets [7]. These basically encode the entities IFCPROJECTEDCRS and IFCMAPCONVERSION as two IFCPROPERTYSET entities named EPSET_PROJECTEDCRS and EPSET_MAPCONVERSION, respectively, attached to the top node of IFC exchange. Their simple attributes are encoded as individual IFCPROPERTYSINGLEVALUE included in the corresponding property set. Of course, all attributes which point to another entity are not included in the property sets (e.g. IFCCOORDINATEOPERATION.TARGETCRS).

For example, consider CRS named *DB_REF2016 zone 4* with EPSG code 9934. This would be serialized as IFCCOORDINATEREFERENCESYSTEM.NAME = 'EPSG:9934' in IFC4 and IFC4.3. In IFC2x3, this would be serialized as a property IFCPROPERTYSINGLEVALUE with .NAME = 'NAME' and .NOMINALVALUE = 'EPSG:9934' residing within a property set IFCPROPERTYSET.NAME = 'EPSET_PROJECTEDCRS' attached to the IFCPROJECT.

3 Proposed Solution

The goal of our solution is to upgrade the level of support for both older versions of IFC when compared to IFC4.3. Our proposal builds upon the idea from Section 2.3: let us introduce more specialized property sets. We introduce three new property sets: EPSET_GEOGRAPHICCRS, EPSET_RIGIDOPERATION and EPSET_MAPCONVERSIONSCALED containing their corresponding attributes as properties. As such, each scenario from Table 1 is enabled and full support for georeferencing scenarios is ensured for IFC2x3.

Because the IFC4 version already supports some of the scenarios, the solution is not as trivial as for IFC2x3. Here, attaching a property set EPSET_MAPCONVERSION to IFCPROJECT would neglect the semantically preferred serialization with IFCMAPCONVERSION. Thus, increased carefulness in workaround design is required. The most suitable solution is to either use only semantically defined entities or the workaround property sets. As such, only one option from Table 1 *topocentric on projected* does not use a workaround, employing IFCMAPCONVERSION and IFCPROJECTEDCRS entities. All other scenarios use the workaround from IFC2x3 with property sets.

Additionally, handling of WKT has to be addressed in a sustainable manner. A WKT is a string and can be easily stored in any attribute of type IFCLABEL or IFCTEXT. The IFCCOORDINATEREFERENCESYSTEM.DESCRIPTION attribute presents itself as the perfect candidate in IFC4 to fulfil this role. The IFCCOORDINATEREFERENCESYSTEM.NAME attribute shall be 'WKT' in this case (as standardized by IFC4.3). This workaround shall be applied appropriately for the custom property sets in IFC2x3 from Section 2.3.

Since the options of georeferencing in IFC4.3 are mutually exclusive already per schema definition (i.e. only one IFCCOORDINATEREFERENCESYSTEM per geometric context), we need address this with strict plausibility rules for the workarounds as well. On the one hand, only one of the property sets representing a CRS and only one of the property sets representing a CO can be attached to the IFCPROJECT, allowing only for combinations from Table 1. On the other hand, it is not allowed to use IFCPROJECTEDCRS and EPSET_PROJECTEDCRS simultaneously. If both are present (for whatever

Table 2: Project base point and CRS's name and EPSG code for 'Torte' examples. Both projected CRS (first two lines) are combined with the vertical CRS *DHHN2016* with EPSG code 7837, while the geographic CRS uses ellipsoidal height.

EPSG	CRS	East	North	Height	code
31467	DHDN / GK Zone 3	3 458 716.16 m	5 439 968.70 m	0 m	$\mod 12 = 0$
25832	ETRS89 / UTM 32N	458 657.30 m	5 438 232.20 m	0 m	$\mod 12 = 1$
4326	WGS84	8.433 679 8°	49.095 550 6°	47.840 m	$\mod 4 = 0$

reason), the IFCPROJECTEDCRS takes precedence and EPSET_PROJECTEDCRS should be ignored. All these rules can be enforced by appropriately expanding the GeoMVD [4]. Their exact serialization would surpass the size limit of the paper and thus remains an exercise for the reader.

4 Proof of Concept

4.1 Example Files: Torte

We provide a series of IFC example files to showcase the correct usages of the standards together with the proposed workarounds. The basis data for our considerations is the example file E.7.1 from the official IFC documentation [1]. It presents a model located close to Karlsruhe, Germany, in three different CRS with their corresponding coordinates for the same project base point as listed in Table 2. In order to cover all possibilities, we devise a split model, where individual pieces in separate IFC files constitute for the whole cake – 'Torte'. This results in 24 files i.e. an IFC file for each combination of the three IFC versions ($code/4 \mod 3$) with the georeferencing scenarios from Table 1 ($code \mod 4$).

4.2 Implementation: KITModelViewer and Squirrel

KITModelViewer [12] is a tool for the visualization and analysis of BIM and geographic information system (GIS) data. Different data models can be displayed in a common, georeferenced environment. For a simple check of the georeferencing of an IFC dataset, there is an extra query dialogue that provides an overview of all georeferencing options in an IFC dataset from [6]. Figure 2 shows the 'Torte' example in the KITModelViewer, where a map is placed underneath for easy visual verification of the real location.

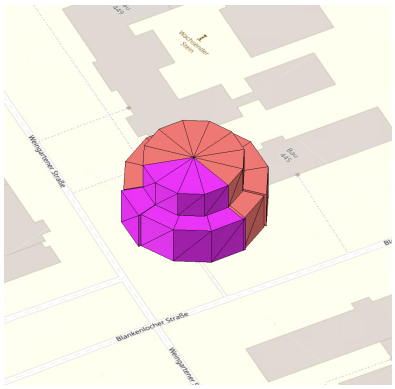


Figure 2: Torte loaded in KIT Model Viewer.

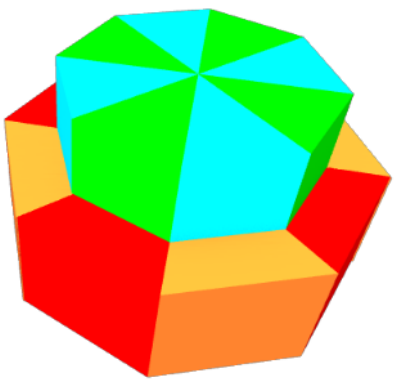


Figure 3: Torte loaded in Squirrel with the proposed property sets EPSET_PROJECTEDCRS and EPSET_RIGIDOPERATION.

Squirrel 3.0 is a comprehensive Common Data Environment (CDE) with a strong focus on openBIM principles. It implements key standards provided by buildingSMART International (bSI), including support for the IFC schema and BIM Collaboration Format (BCF)-based communication, both via file exchange and programming interfaces. Squirrel is primarily applied in large-scale infrastructure projects, where experience reveals a wide range of challenges related to CRS, particularly due to the high number of partial models involved. It is not uncommon to visualize and coordinate hundreds of IFC models within a single project. When railway infrastructure and station buildings are combined, at least one stakeholder often needs to adopt a foreign CRS to ensure interoperability across the project. Figure 3 shows the 'Torte' example in the Squirrel CDE, with the proposed property sets seen on the right-hand side.

5 Conclusions

Last year, a new version of the IFC standard was published – so called IFC4.3 – which adds new entities and concepts to support various georeferencing scenarios as shown in Figure 1 [1], [3]. While this enables crisp data exchanges in the future, the question arises about the previous official versions of IFC: can these be updated in a non-intrusive manner to support the scenarios as well?

This contribution proposes an extension in the form of specialized property sets based on [7], which are attached to the top node of the IFC data structure as described in Section 3. We have limited ourselves to reuse existing mechanisms in IFC standard in order to remain compatible with existing implementations. All-in-all, 5 property sets for each of the 5 IFC entities from Table 1 have been introduced. Thus, we have completed the possibilities for georeferencing across all three currently official IFC versions without the need for a new version of the established standards.

The proposed solution was prototypically tested with a series of carefully designed example files for all possible combinations of IFC versions and georeferencing scenarios for two different projected and a geographic CRS as described in Section 4. Two independent software solutions – KITModelViewer and Squirrel – have implemented the proposed workaround and achieved correct interpretation of each combination as seen in Figures 2 and 3. This assures the solution as viable and we invite all other software solutions to support these workaround as well.

The solution is not perfect, though. The property sets simulating CO connecting the CRS to the 3D geometric representation is not as explicit as in the file exchanges of the newer versions, with explicit attributes SOURCECRS and TARGETCRS on entity IFCCOORDINATEOPERATION. There is no possible way to explicitly include such information in a property, thus leaving only the documentation to enforce that such georeferencing applies to the 3D geometric context. Let it be noted on this place, that [7] does not cover this implicit definition explicitly in their documentation either.

5.1 Future Works

While studying the documentation and discussing the possible workaround strategies, the authors noted that the placement concept (IFCOBJECTPLACEMENT and derived classes) provide no explicitly serialized information about the geometric context they reside within. Thus, there is no mechanism in place which would allow to interpret the distances in placement in correct context. This should be addressed in the next versions of IFC.

Additionally, the example files do not account for a change in geometry due to the different distortions as a consequence of the chosen CRS and base point location. The series shall be updated and/or expanded to account for these peculiarities as well.

References

[1] International Organization for Standardization, “ISO 16739: Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries – Part 1: Data schema”, Geneva, CH, Standard, 2024.

[2] Š. Jaud, C. Clemen, S. Muhič, and A. Borrmann, “Georeferencing in IFC: Meeting the Requirements of Infrastructure and Building Industries”, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. X-4/W2-2022, pp. 145–152, 2022. DOI: 10.5194/isprs-annals-X-4-W2-2022-145-2022

[3] Š. Jaud, C. Clemen, L. Wikström, S. Muhič, and A. Borrmann, “Georeferencing of IFC Geometries”, in *Proc. of the 20. Internationaler Ingenieurvermessungskurs*, Zurich, CH, 2023.

[4] Š. Jaud and C. Clemen, “GeoMVD: the Journey to High-Quality Georeferencing Profiles in IFC Datasets”, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. X-4/W5-2024, pp. 203–210, 2024. DOI: 10.5194/isprs-annals-X-4-W5-2024-203-2024

[5] buildingSMART Deutschland, *Sechs Softwarehäuser absolvieren BIM Fit Check Georeferenzierung 2025 erfolgreich*, 2025. [Online]. Available: https://www.buildingsmart.de/buildingsmart/aktuelles/sechs-softwarehaeuser-absolvieren-bim-fit-check-georeferenzierung-2025

[6] C. Clemen and H. Görne, “Level of Georeferencing (LoGeoRef) using IFC for BIM”, *Journal of Geodesy, Cartography and Cadastre*, vol. 2019, no. 10, pp. 15–20, 2019.

[7] J. Mitchell et al., “User Guide for Geo-referencing in IFC”, buildingSMART International, Tech. Rep., 2020. [Online]. Available: https://www.buildingsmart.org/wp-content/uploads/2020/02/User-Guide-for-Geo-referencing-in-IFC-v2.0.pdf

[8] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, “A design science research methodology for information systems research”, *Journal of management information systems*, vol. 24, no. 3, pp. 45–77, 2007.

[9] International Organization for Standardization, “ISO 19111: Geographic information – Referencing by coordinates”, Geneva, CH, Standard, 2019.

[10] International Association of Oil & Gas Producers, *EPSG Geodetic Parameter Dataset*, 2021. [Online]. Available: https://epsg.org/

[11] M. Rives et al., *IFC-Tunnel Project: Requirement Analysis Report*, 2020. [Online]. Available: https://app.box.com/s/uwb9yngfclixts81coeqg4ncl6fs832j

[12] A. Geiger, *SDM_KITModelViewer*, 2025. [Online]. Available: https://github.com/KIT-IAI/SDM_KITModelViewer

Integriertes und skalierbares Planungssystem auf Basis von GIS- und Bodeninformationen in der Infrastrukturplanung

Maik Görgner[1,2], Simon Hoeng[1], Anna-Maria Mehringer[1,2] and Jonas Wiederer[1]
[1]Ostbayerische Technische Hochschule Regensburg, Seybothstraße 2, 93053 Regensburg
[2]Firmengruppe Max Bögl, Max-Bögl-Straße 1, 92369 Sengenthal
E-mail(s): maik.goergner@oth-regensburg.de, simon.hoeng@oth-regensburg.de,
anna-maria.mehringer@oth-regensburg.de, jonas.wiederer@oth-regensburg.de

Abstract: Die Energiewende erfordert den beschleunigten Ausbau des deutschen Stromnetzes und der dazugehörigen Infrastruktur. Herausforderungen wie heterogene Umweltbedingungen, vielfältige Hindernisse und wechselhafte Bodenverhältnisse machen dieses Vorhaben zu einer hochkomplexen Planungsaufgabe. Die effiziente und nachhaltige Abwicklung dieser Energieinfrastrukturprojekte erfordert daher innovative und integrative Lösungsansätze.

Vor diesem Hintergrund schlägt die Arbeit ein ganzheitliches, innovatives System vor, das geotechnische, umweltbezogene und infrastrukturelle Daten modular miteinander verknüpft. Ziel ist eine automatisierte, kontextabhängige Auswahl und Kombination geeigneter Bauverfahren innerhalb linearer Infrastrukturprojekte. Der geometrische Verlauf dieser Projekte wird importiert und eine automatisierte Querungserkennung identifiziert Straßen-, Fluss- und Waldpassagen. Unter Berücksichtigung des Ziels eines möglichst konsistenten Bauverfahrenseinsatz unterteilt das System die Route in standardisierte Module, die jeweils durch ein spezifisches Bauverfahren und Bodencharakteristika definiert sind. Jedes Modul enthält realistische Kennzahlen zu Ressourcenaufwand, Bauzeit und Kosten sowie variable Parameter.

Innovativ ist dabei die systematische Zusammenführung heterogener Datensätze aus Geotechnik sowie Umwelt- und Infrastrukturinformationen zu einem integrierten Bewertungsrahmen. Gegenstand der Forschung ist die optimale Modellierung und algorithmische Verknüpfung multidimensionaler Daten, um ein effektives und effizientes Tool zu generieren. Zusätzlich untersucht die Arbeit mögliche Validierungsstrategien (Abgleich mit echten Bauprojekten, Sensitivitätsanalysen) sowie Performance-Metriken (Rechenaufwand, Genauigkeit der Empfehlung) für das modulare Entscheidungssystem.

Keywords: GIS, BIM, Geotechnics, Modularisation, Optimisation

DOI: 10.18154/RWTH-CONV-254901. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Motivation, Konflikt und Herausforderungen

Der beschleunigte Ausbau des deutschen Stromnetzes gilt als eine der zentralen Voraussetzungen für das Gelingen der Energiewende. Insbesondere großflächige Linienbaumaßnahmen stellen Planerinnen und Planer vor komplexe Herausforderungen, die über klassische Ingenieurleistungen hinausreichen. Im Fokus steht nicht nur die enorme Länge dieser Infrastrukturachsen, sondern vor allem ihre Durchdringung hochdiverser Raumtypen, von urbanen Zonen bis hin zu ökologisch sensiblen Naturräumen [1]. Daraus ergibt sich ein Spannungsfeld zwischen Zeitdruck, ökologischen Zielsetzungen und bautechnischen Rahmenbedingungen, die im Rahmen bestehender Verfahren parallel berücksichtigt werden müssen [2]. Diese hohen Anforderungen an eine präzise und gleichzeitig effiziente Trassenplanung führen zu einem wesentlichen Zielkonflikt im Zuge der Projektabwicklung. Umweltbedingungen, Schutzgebiete, bestehende Infrastrukturen sowie geotechnische Unsicherheiten erfordern eine Vielzahl standortspezifischer Abwägungen [3]. Konventionelle Planungsprozesse basieren bislang auf einer Kombination aus manuell aufbereiteten Fachgutachten, projektbezogener Expertise und iterativen Abstimmungen mit Beteiligten. Diese Praxis ist jedoch zeit- und ressourcenintensiv und lässt sich nur schwer skalieren oder digital automatisieren. Besonders herausfordernd ist die systematische Auswahl geeigneter Ausführungsvarianten entlang der geplanten Trasse.

Die Entscheidung für ein ungeeignetes Verfahren kann nicht nur zu Kostensteigerungen und Verzögerungen führen, sondern auch Genehmigungsrisiken erhöhen oder ökologische Zielsetzungen konterkarieren. Vor diesem Hintergrund gewinnt der Einsatz datenbasierter, integrativer Planungsansätze zunehmend an Bedeutung. Zum einen sind Werkzeuge erforderlich, welche geotechnische, infrastrukturelle und ökologische Parameter erfassen und aggregieren. Zum anderen müssen diese Parameter innerhalb einer systematischen und automatisierbaren Entscheidungslogik bewertet werden. Nur so lässt sich der gestiegene Komplexitätsgrad der Trassen- und Bauverfahrenplanung in einer für Großprojekte skalierbaren Form bewältigen.

Speziell die Wahl der jeweiligen Bauweise stellt einen bislang wenig standardisierten, stark erfahrungsbasierten Schritt dar. Diesen Schritt zu optimieren, könnte jedoch maßgeblich zur Risikominimierung, Kostensicherheit und Genehmigungsfähigkeit beitragen. In Anlehnung an [4] eröffnen datenbasierte sowie modellgestützte Vorgehen hier das Potenzial, planungsrelevante Faktoren frühzeitig transparent zu machen. Auch können Varianten systematisch bewertet und robuste, nachvollziehbare Entscheidungen abgeleitet werden. Die hier beschriebene Methodik setzt genau an dieser Schnittstelle an. Ziel ist es, ein methodisches Fundament für ein Planungsinstrument zu entwickeln, das den Bauverfahrensentscheid entlang einer Trasse quantitativ unterstützt.

2 State of the Art und Forschungsbedarf

2.1 Grundlagen aktueller GIS- und Geodatenintegration

Die Nutzung von Geoinformationssystemen (GIS) hat sich in den vergangenen Jahren als elementarer Bestandteil der Infrastrukturplanung etabliert. Insbesondere für Trassenbauprojekte liefern standardisierte Geodatenformate und Geo-Web-Services (z. B. WMS, WFS nach OGC) die technischen Voraussetzungen für die Zusammenführung heterogener räumlicher Informationen.

[5] zeigen am Beispiel des Projekts „3D Tracks“ bereits früh, wie interoperable Geo-Web-Services die Planungszusammenarbeit verbessern und aktuelle Geodaten in die Trassenplanung integrieren können. Diese grundlegende Fähigkeit zur Datenintegration liefert zugleich die technische Infrastruktur für den hier verfolgten Ansatz.

Im Leitfaden 3D-GIS und Energie (2015) unterstreichen [6] das wachsende Potenzial von 3D-GIS-Systemen in der Energieinfrastrukturplanung, insbesondere im Hinblick auf Visualisierung, Datenintegration und präzisere geometrische Analysen. Hervorzuheben ist dabei die Diskussion raumwiderstandsbasierter Verfahren, wie sie etwa durch Least Cost Path Analysis oder Multi-Criteria Decision Analysis umgesetzt werden. Beide Methoden basieren auf der Rasterung raumbezogener Einflussfaktoren. Kriterien wie Bebauung, Artenschutz oder ökologische Einschränkungen können über Zellwiderstände berücksichtigt werden. Die Auswahl geeigneter Trassenabschnitte erfolgt auf Grundlage gewichteter Entscheidungsregeln.

Auch kommerzielle Lösungen wie nisStrom Trasse nutzen GIS-basierte Systeme zur Planung und Verwaltung von Stromtrassen. Sie fokussieren sich jedoch primär auf die Verwaltung, Dokumentation und geometrische Darstellung von Trassenverläufen, ohne eine automatisierte Ableitung optimaler Ausführungsvarianten vorzusehen. [7]

Auch im Bereich der digitalen Bauwerksmodellierung zeigt sich eine zunehmende Integration geodätischer Methoden in moderne Planungsprozesse.

In diesem Zusammenhang stellt [8] im Leitfaden Geodäsie und BIM Version 4.0 (2024) fest, dass klassische BIM-Systeme bislang nur eingeschränkt für langgestreckte Infrastrukturprojekte geeignet sind. Das verdeutlicht die Wichtigkeit, bestehende Fachsysteme gezielt mit digitalen Modellierungsansätzen zu verknüpfen. Besonders bei komplexen räumlichen Strukturen und technischen Randbedingungen zeigt sich der Bedarf an integrierten, datenbasierten Lösungsansätzen. [8] beschreibt hierzu einen hybriden Ansatz, bei dem Trassierungssoftware eingesetzt wird und die resultierenden Planungsdaten über Austauschformate in BIM-konforme 3D-Modelle überführt werden.

Eine internationale Untersuchung von [9] bestätigt diesen Trend hin zu integrierten, modellbasierten Planungsansätzen. Die Autoren betonen einerseits die Vorteile solcher Systeme in Bezug auf Genauigkeit, Effizienz und Koordination. Andererseits wird auf bestehende Interoperabilitätsprobleme und die Notwendigkeit standardisierter Verfahren verwiesen.

Eine ähnliche Richtung verfolgt [10]. Anhand eines realen Infrastrukturprojekts (Stadtautobahn M-30 in Madrid) wird in diesem Artikel dargelegt, wie sich BIM-Modelle über Formate wie IFC und SLPK in GIS-Anwendungen überführen lassen. Die Anbindung an eine externe Datenbank ermöglicht dabei Aktualisierungen, ohne das 3D-Modell selbst verändern zu müssen. Echtzeitinformationen können somit leichter in das gesamte Modell integriert werden.

Besonders hervorzuheben sind auch die Entwicklungen aus dem maschinellen Tunnelbau. So entwickelte das Team um [4] im Rahmen des Sonderforschungsbereichs 837 an der Ruhr-Universität Bochum ein interaktives Entscheidungssystem, das auf Basis simulationsgestützter Vortriebsmodelle Tunnelvarianten hinsichtlich Bauzeit, Kosten und Risiken bewertet. Hierbei werden Freiformtrassen erzeugt, simulationsbasiert bewertet und Varianten in Echtzeit verglichen.

Diese Methodik zeigt exemplarisch das Potenzial regelbasierter Entscheidungsplattformen in hochkomplexen Planungssituationen.

2.2 Abgrenzung und Beitrag der vorliegenden Arbeit

Die vorliegende Arbeit greift diese raumanalytischen Vorarbeiten auf, geht jedoch über bestehende Modelle hinaus. Während frühere Arbeiten primär den optimalen Trassenverlauf adressieren, fokussiert der hier entwickelte Ansatz auf die algorithmische Ableitung geeigneter Bauverfahren entlang bereits festgelegter Trassen. Die Verknüpfung geotechnischer, umweltbezogener und infrastruktureller Einflussgrößen auf modularer Ebene sowie deren Integration in ein automatisiertes Bewertungssystem stellt somit eine inhaltliche und methodische Weiterentwicklung bestehender GIS-gestützter Planungsansätze dar. Auch erweitert die in dieser Arbeit entwickelte Methodik die oben genannten Ansätze um eine prozedurale Entscheidungslogik. Basierend auf Querungsmerkmalen und Bodendaten werden konkrete Bauweisen abgeleitet und systematisch in standardisierte Module überführt. Damit wird aus einer klassischen Trassenbewertung ein durchgängiges, verfahrensspezifisches Planungssystem abgeleitet. Zahlreiche Forschungsarbeiten leisten demnach wertvolle Beiträge in Bereichen wie Geodatenintegration, 3D-GIS, Raumwiderstandsanalyse, Trassendokumentation und Variantenbewertung. Integrative Verfahren zur automatisierten Bauverfahrenswahl entlang vorgegebener Geometrien hingegen stellen bislang einen weitgehend unerforschten Ansatz dar. Die hier beschriebene Forschungsarbeit zielt darauf ab, diese Lücke zu schließen.

Durch die Entwicklung eines datengetriebenen, modularen Entscheidungssystems sollen heterogene Geodatenquellen algorithmisch verknüpft und systematisch ausgewertet werden. Im Gegensatz zu bisherigen Planungsansätzen wird die Ausführungsvariante hier nicht mehr separat durch Expertenentscheidungen getroffen, sondern auf regelbasierte Kriterien gestützt und konsistent entlang der gesamten Trasse automatisiert durchgeführt. Dadurch entsteht eine skalierbare, transparente und für Variantenvergleiche unmittelbar nutzbare Planungslogik, die sowohl wissenschaftlich neuartig als auch praktisch anwendbar ist.

3 Methode und Konzeptidee

3.1 Grundidee des Modulsystems

Zentrales Element des vorgestellten Konzepts ist die Unterteilung einer festgelegten Linienbaustelle in standardisierte Module. Jedes dieser Module wird durch das zugewiesene Bauverfahren, die maßgebliche Bodenklasse sowie variable Parameter wie die Abschnittslänge eindeutig charakterisiert. In der Folge ermöglicht die modulare Gliederung eine standardisierte Bewertung technischer Optionen und eine differenzierte Ressourcenplanung sowie fundierte Variantenvergleiche. Die methodische Umsetzung basiert auf einer Vielzahl digitaler Datengrundlagen. Dazu gehören insbesondere der digitalisierte Trassenverlauf (z.B. im KML- oder SHP-Format), vektorbasierte Infrastrukturinformationen wie Straßen, Flüsse oder Waldflächen sowie geotechnische Klassifikationen. Diese Informationen bilden den Bezugsrahmen für eine automatisierte räumliche Analyse sowie die Bewertung des Trassenumfelds.

3.2 Methodische Umsetzung und Datenstruktur

Das methodische Zielbild der Forschungsarbeit sieht dabei die in Abb. 1 dargestellte zweistufige Prozesskette vor, in der die vorgestellte Systematik praxisorientiert umgesetzt wird.

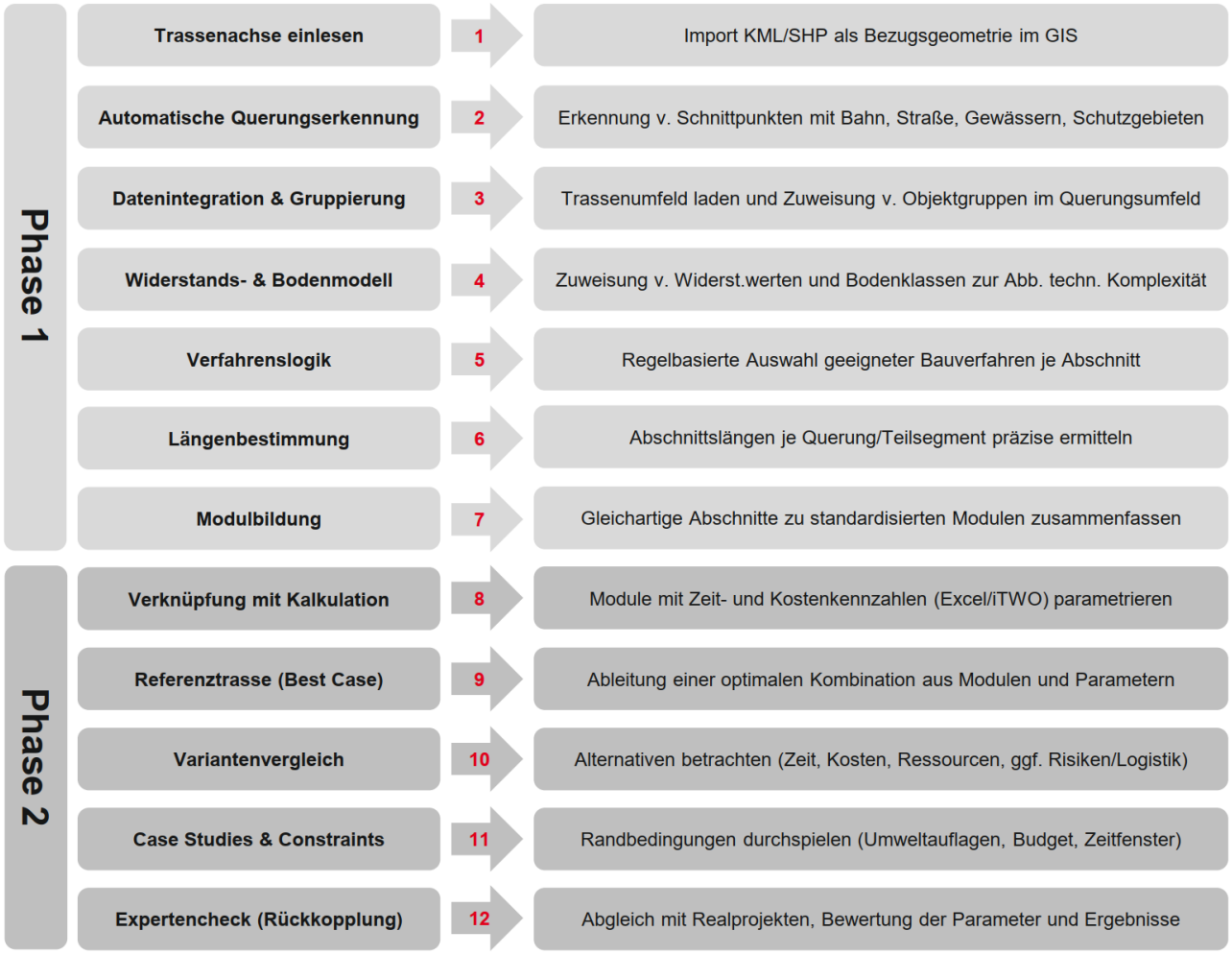


Abbildung 1: Gesamtprozess der Forschungsarbeit – Modulare Trassenplanung in zwei Phasen

Zentrales Element der methodischen Umsetzung ist die Entwicklung einer strukturierten, regelbasierten Datenbankarchitektur. Hierbei werden geographische, technische und wirtschaftliche Parameter entlang einer Trasse systematisch miteinander verknüpft. Jeder Abschnitt der Linienbaustelle wird als standardisiertes Planungselement (Modul) beschrieben. Die spezifischen Eigenschaften eines Moduls ergeben sich aus der Kombination von Querungstyp, Bodeninformation und variablen Parametern. Die zugrundeliegenden Eingangsdaten umfassen die Liniengeometrie, vektorisierte Infrastruktur- und Umweltdaten sowie bodenkundliche Informationen.

Diese werden automatisiert mit der Linienbaustelle verschnitten, um relevante Querungspunkte zu erkennen und geometrisch zu quantifizieren. Im Zuge einer regelbasierten Analyse wird jeder Querung auf Basis von Länge und Bodenklasse ein geeignetes technisches Verfahren zugewiesen. Die dabei zur Anwendung kommende Entscheidungslogik beruht auf projektinternem Expertenwissen, das in Form von Schwellenwerten und Bedingungen in die Datenstruktur integriert ist. Die so festgelegten

Bauverfahren sind wiederum mit hinterlegten Kennwerten zu Bauzeit, Ressourceneinsatz und Kosten verknüpft. Somit ergibt sich eine technische und wirtschaftliche Parametrisierung aller vorkommenden Module.

Die resultierenden Module werden in einer einheitlich aufgebauten Datenstruktur gespeichert. Mithin fungiert diese innovative Datenstruktur als zentrales Bindeglied zwischen der raumbezogenen Analyse im GIS und der betriebswirtschaftlichen Weiterverarbeitung.

Sie erlaubt es, Trasseninformationen automatisiert in planungsrelevante Entscheidungsstrukturen zu überführen und bildet zugleich den Ausgangspunkt für Variantenvergleiche und eine spätere Integration in digitale Planungsumgebungen. Der Mehrwert liegt in der strukturierten Formalisierung bislang weitgehend erfahrungsgeleiteter Entscheidungsprozesse.

4 Erste prototypische Umsetzung als methodischer Proof-of-Concept

Im Rahmen eines ersten prototypischen Ansatzes wurde die automatisierte Erkennung und Klassifizierung trassenrelevanter Querungen in der QGIS erfolgreich realisiert. Als methodischer Ausgangspunkt diente der Verlauf einer Linienbaustelle, der als Shape-Datei (.shp) in das System eingelesen wurde. Darauf aufbauend wurde ein Prozess etabliert, der die systematische Überlagerung der Trasse mit thematischen Geodatenlayern (z.B. Straßen und Gewässern) und die Identifikation aller relevanten Schnittpunkte ermöglicht. Diese Querungen werden in Abb. 2 dargestellt.

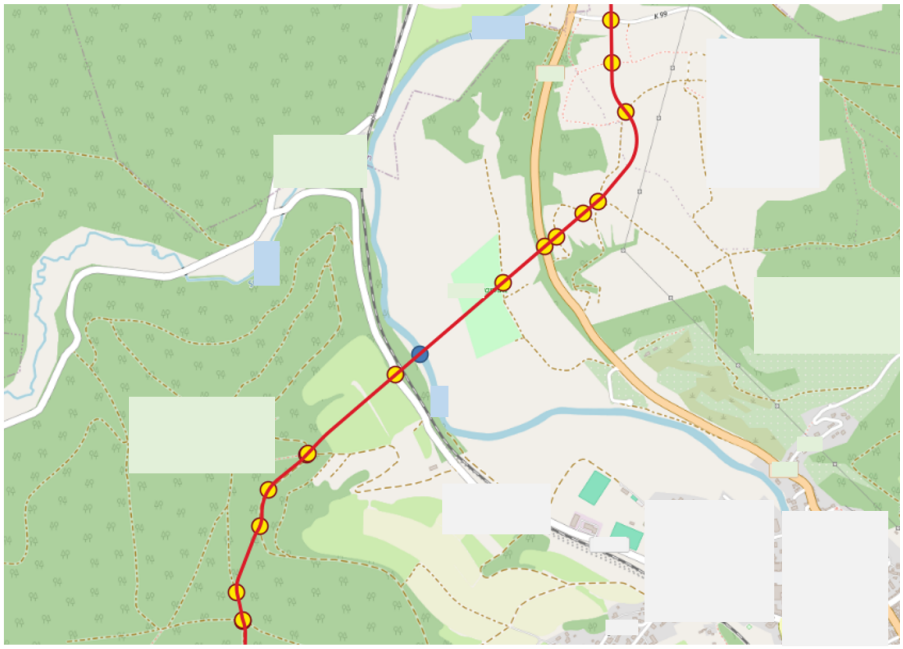


Abbildung 2: Automatisierte Querungserkennung in QGIS entlang der Trassenachse

Die automatisierte Auswertung erfolgt über regelbasierte Abfragen im QGIS-Modellbuilder. Diese extrahieren jede Schnittstelle als eigene Geometrie, versehen sie mit Attributen und überführen sie in eine strukturierte Tabelle mit Lage-, Typ- und Geometriedaten.

Anknüpfend an diese Einteilung können beispielsweise Cluster von Querungen oder besonders konfliktbelastete Bereiche identifiziert werden. Zudem ermöglicht das System die automatisierte

Analyse zusätzlicher Geodatensätze wie Boden-, Schutz- oder Katasterdaten zur gezielten Bewertung potenzieller Planungseinschränkungen. Die prototypische Umsetzung belegt die Machbarkeit einer teilautomatisierten Trassenanalyse und verdeutlicht das Potenzial zur strukturierten, effizienteren Gestaltung raumbezogener Planungsprozesse.

5 Ausblick, zukünftige Arbeit und Contribution

Mit der Entwicklung eines funktionalen Prototyps wurde ein erster zentraler Meilenstein auf dem Weg zu einer automatisierten, datenbasierten Trassenplanung erreicht. In den nächsten Entwicklungsschritten wird insbesondere daran gearbeitet, geotechnische Datenquellen so zu integrieren, dass daraus bodenspezifische Kennwerte für die Modulbildung abgeleitet werden können. Parallel dazu wird die Integration vertiefter bodenmechanischer Informationen vorangetrieben, insbesondere durch die Einbindung standardisierter Bodenklassifikationen und Tiefeninformationen aus digitalen Erkundungsdaten.

Ein nächster Schritt besteht in der vollständigen Automatisierung der Modulbildung sowie der algorithmischen Zuweisung der Bauweise auf Grundlage technischer, geotechnischer und umweltbezogener Eingangsparameter. Gleichzeitig soll die Auswahl geeigneter Module künftig automatisiert erfolgen, basierend auf einem klar definierten Regelwerk, das raumbezogene Bedingungen berücksichtigt. Ergänzend dazu wird angestrebt, erste Variantenvergleiche durchzuführen, um unterschiedliche Trassenlösungen im Hinblick auf Ressourceneinsatz, Bauzeit und Kosten bewerten zu können.

Zur technischen Absicherung des Modells erfolgt eine systematische Validierung anhand realer Trassenverläufe und Projektkontexte. Im Fokus steht hierbei die Frage, inwieweit das System verlässlich auf unterschiedliche Ausgangsbedingungen reagiert. etwa hinsichtlich Trassenlänge, möglicher Bauweisen oder der Komplexität der Umgebung. Da im produktiven Einsatz stets mit geometrisch definierten Trassenverläufen gearbeitet wird, muss das System in der Lage sein, eine solche Eingabe vollständig automatisiert zu verarbeiten. Dies umfasst die Einlesung, die Modulgenerierung sowie die Übergabe an Folgeprozesse wie Mengenermittlung oder Kalkulation.

Im Rahmen dieser Arbeit soll ein stabiles, integratives Planungssystem entstehen, das auch bei wachsendem Datenvolumen zuverlässig arbeitet und nahtlos in bestehende digitale Arbeitsumgebungen eingebunden werden kann. Historische Entscheidungs- und Ausführungsdaten sollen genutzt werden, um algorithmische Vorschläge zur Bauverfahrenswahl kontinuierlich zu verbessern und kontextspezifisch zu verfeinern. Durch die geplante Weiterentwicklung entsteht ein praxisnahes Werkzeug, das Fachplaner bereits in frühen Projektphasen gezielt unterstützt und den Entscheidungsprozess strukturiert. Auf diese Weise kann ein konkreter Beitrag zu einer beschleunigten, effizienteren und nachhaltigeren Umsetzung von Linienbaustellen geleistet werden.

Literatur

[1] TenneT TSO GmbH. »Tennet - Projekt SuedLink«. Adresse: https://www.tennet.eu/de/projekte/suedlink

[2] P. Ahmels und E. Bruns, »Auswirkungen verschiedener Erdkabelsysteme auf Natur und Landschaft (EKNA FKZ 3514 82 1600)«, Deutsche Umwelthilfe, Berlin, 2016, S. 12–24, 86–98.

[3] E. Neuling, »Stromfluss unter der Erde – Einsatz von Erdkabeln beim Übertragungsnetzausbau«, Naturschutzbund Deutschland (NABU) e. V., Berlin, 2013, S. 1–6.

[4] M. König u. a., »Prozessorientierte numerische Vortriebsmodellierung – Trassenplanung für den innerstädtischen Tunnelbau«, *tunnel - Fachzeitschrift für Tunnelbau und Infrastruktur*, S. 48–53, Juni 2021.

[5] H. Steuer und A. Donaubauer, »Geo-Web-Services im Rahmen der kollaborativen Trassenplanung«, in *Geoinformationssysteme - Beiträge zum 17. Münchener Fortbildungsseminar*, Berlin und Offenbach: Herbert Wichmann Verlag, 2012, S. 159–167.

[6] J. Schito und S. Grassi, »Berechnung optimaler HSL-Pfade und Korridore für Hochspannungsleitung«, in *Leitfaden 3D-GIS und Energie*, München: Runder Tisch GIS e. V. und TUM, 2015, S. 80–82.

[7] T. Hermes. »nisStrom Trasse«, GIS Consult GmbH - Gesellschaft für angewandte geographische Informationssysteme. Adresse: https://www.gis-consult.de/produkte-und-loesungen/fachschale-nisstromtrasse/

[8] A. Borrmann, »Building Information Modeling für den Infrastrukturbau«, in *Leitfaden Geodäsie und BIM - Version 4.0*, Bühl und München: DVW e. V. und Runder Tisch GIS e. V., 2024, S. 42–54.

[9] M. Arashpour u. a., »Enhancing infrastructure planning and design through BIM-GIS integration«, *Structure and Infrastructure Engineering*, S. 1–19, Feb. 2025.

[10] J. Cepa u. a., »Integrating BIM and GIS for an Existing Infrastructure«, *Applied Sciences (MDPI)*, S. 1–17, Nov. 2024.

TopoBEM: Using Non-Manifold Topology and Building Energy Simulation to Facilitate Evidence-Based Decision Making in Architectural Design

Malte Junge[1]
[1]Design Computation (DC), Faculty of Architecture, RWTH Aachen University, Aachen, Germany
E-mail(s): junge@dc.rwth-aachen.de

Abstract: Building Energy Modelling (BEM) and energy simulation offer great potential for evidence-based decision-making during the early stages of building design. However, the models and workflows of established Building Information Modelling (BIM) programs are not ideal for energy simulation and early phase design exploration. This paper explores *Non-Manifold Topology* (NMT) as an innovative computational method for digital building models, that produces lightweight, adaptable models that are well-suited for energy simulation. NMT models differentiate themselves from BIM models, by explicitly representing rooms as discrete topologic cells, rather than describing them implicitly through an envelope of physical building elements. The main contribution is an integrated workflow, that generates conceptual room-cell models enriched with semantic data and directly transferable to established building performance simulation tools. By applying this workflow to a case study building, this research demonstrates how NMT can bridge the gap between conceptual design and BEM.

Keywords: building energy modelling, non-manifold topology, evidence-based design, thermal comfort, daylighting

DOI: 10.18154/RWTH-CONV-254873. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

In Building Information Modelling (BIM) buildings are represented as spatial assemblies of semantically and geometrically rich digital assets. This approach has transformed the way planners represent early phase designs digitally, by prompting them to make relatively detailed assumptions about scale, construction, and material early on in the design process. This early commitment may limit creative exploration and lead to time-consuming and hardware-intensive workflows during the rapidly evolving conceptual phase. Consequently, many architects view BIM authoring tools as suboptimal for conceptual planning, the phase where the most impactful decisions affecting cost, schedule, architectural language, and environmental impact are made [1]. This is critical, because even though conceptual designs are often perceived as flexible, research shows that initial design decisions tend to persist throughout a project [2].

Although BIM and design computation programs offer powerful, evidence-based tools, such as quantity calculations, structural analysis, and Building Energy Modelling (BEM), to support these decisions, their integration into early design remains fragmented due to usability issues and limited interoperability. Often, designers have to convert or remodel models between applications, which disrupts the design process and reduces overall productivity [3]. Thus, integrated workflows between BIM and BEM have become increasingly important in the AEC sector. Designers who invest significant effort in detailed BIM models naturally want to use them for energy simulation too, but establishing industry standards for data exchange has proven challenging [4].

To address these challenges, this paper introduces an integrated workflow that harnesses *non-manifold topology* (NMT) for the creation of lightweight and adaptable room-cell models. Unlike BIM representations of physical building elements, NMT-based models emphasize essential spatial and semantic properties on the room-level without overburdening designers with unnecessary detail. When combined with BEM, this approach provides rapid, evidence-based feedback on key performance indicators, such as thermal comfort and daylight distribution, in the early design stages, while retaining full-scale BIM compatibility for the later phases. Such integration not only streamlines digital workflows but also promotes early, informed decision-making, ultimately leading to more sustainable and cost-effective building designs.

2 Background and Related Work

The concept of using non-manifold topology (NMT) for digital building models was introduced by Robert Aish and Aparajit Pratap in 2013, who proposed to “consider architecture not so much as a physical building model, but [...] as the ‘enclosure of space’” [5, p. 1]. NMT allows representations of geometries that standard modelling kernels might consider degenerate, such as closed volumes with internal surfaces. This enables powerful modelling operations and spatial queries that are not possible in conventional tools. The potential of NMT in computer-aided design has been applied in fields ranging from marine [6] to mechanical engineering and even spatial reasoning in the built environment [7]. However, it did not see meaningful usage in architectural design until the advent of NMT room-cell models.

These models use zero-thickness surfaces to partition closed volumes into sub-volumes. NMT’s perfectly matched room-bounding surfaces yield cleaner subsurface divisions for modelling glazing or façade panels [8] and they can serve as a skeletal framework onto which BIM families are hosted, thus producing change-resilient models that remain adaptable from conceptual design through to fabrication planning. Case studies have confirmed NMT models’ compatibility with AEC tools for structural design, fabrication planning and building performance simulation [9].

These research efforts eventually converged into *Topologic*, an in-development open-source NMT-modelling framework for the AEC industry, that offers implementations for a number of different toolkits and authoring software, such as IfcOpenShell, Dynamo and Grasshopper. Topologic leverages the Open Cascade Technology (OCCT) 3D geometry library, because it provides one of the most comprehensive implementations of non-manifold topology across both open-source and proprietary CAD kernels [10]. Topologic structures spatial data hierarchically into eight main classes, `Vertex`,

`Edge`, `Wire`, `Face`, `Shell`, `Cell`, `CellComplex`, and `Cluster`, with each class aggregating entities from the previous level. Any geometry should be modelled as a topology of the highest applicable order. Buildings, being an arrangement of multiple room-cells, are typically modelled as a CellComplex, or a cluster of multiple CellComplexes. Additional core classes like `Content`, `Context` and `Dictionary` enable semantic enrichment while the `Graph` class provides topological linking, such as connecting adjacent room cells. Lastly, the `Aperture` class is used to incorporate openings and subsurfaces like doors and windows into other Topologies.

3 Method

To evaluate the potential of NMT-models and BEM for evidence-based conceptual design, an end-to-end workflow, from model creation to result evaluation, was developed, using Topologic and a realistic test case. An ideal test case features a clear typology and a utilitarian layout with geometrically similar yet functionally diverse rooms. Hospitals were initially considered, but smaller medical typologies (specialized clinics or outpatient centres) were selected due to complexity. A three-storey 'Medical Care Centre' competition proposal in southwestern Germany was chosen as a good match in scale and programme. The proposal's design brief emphasizes the importance of daylight, outside views, and indoor comfort, to create a healing environment; noting necessary trade-offs between daylight access and solar gains to achieve this goal. The proposal was recreated as a Topologic NMT model to evaluate and improve these design aspects in an iterative and evidence-based manner.

At the time of research, a number of different Topologic implementations were available, including a multi-platform Python module [11] and plug-ins for Dynamo (Revit), Grasshopper (Rhino3D), and Sverchok (Blender). Sverchok offers a more advanced version of Topologic, while Grasshopper integrates seamlessly with Ladybug Tools, offering access to the established daylight and energy simulation engines Radiance and EnergyPlus/OpenStudio. Because no single implementation met all requirements, a hybrid workflow was adopted: Blender for reference geometry, Sverchok for NMT/energy model generation, and Grasshopper for simulation, analysis, and visualisation."

3.1 Topologic Model

The reference design can be represented as a single, continuous `CellComplex` in Topologic. Various workflows can generate a `CellComplex`, but model quality varies significantly depending on the approach used. Higher-order topologies are generally constructed from lower-order entities via constructor functions and Boolean operations; a common workflow slices a perimeter `Cell` with interior `Faces` (representing interior walls) to form partitioned rooms.

Creating the perimeter `Cell` and `Faces` manually as mesh or BREP objects and referencing them into the Topologic workflow, grants designers direct control over the placement of cutting planes, enabling them to create NMT-models through conventional 3D-modelling workflows. However, this method risks the introduction of modelling inaccuracies, such as mismatched partition surfaces, which can lead to problems in the energy simulation because they compromise the watertightness of room-cells and generate redundant geometry.

Thorough experimentation yielded a more robust approach: by using the cutting `Faces` to only slice the floor planes of the building (see Figure 1), instead of the whole volume, the slicing is simplified

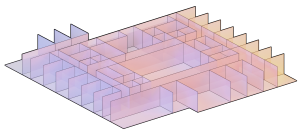


Figure 1: Cutting `Faces` and floor `Face`

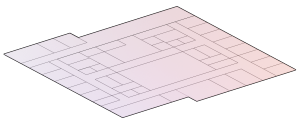


Figure 2: Floor plan `Faces`

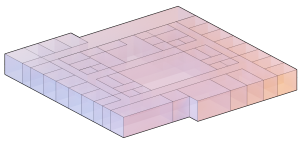


Figure 3: Room `Cells`

to a 2D operation, greatly reducing the potential for degenerate geometry to occur. The result is a schematic plan of each floor, made up of connected `Faces` representing the individual rooms (Figure 2). Extruding these `Faces` in the upwards Z-direction produces watertight, `Cells` with perfectly matched partition surfaces and no superfluous topologic entities, resulting in a room-cell model suitable for simulation (Figure 3).

To enhance the spatial model with semantic information, `Dictionaries` containing room names and types are systematically assigned to all `Cells`. Finally, `Apertures` are added to the model to represent windows and doors.

The semantically enriched `CellComplex` is now ready to be exported into the *HBJSON* format, a serialised energy model schema that provides interoperability with the Ladybug/Honeybee simulation environment in Grasshopper. In the simulation environment, the model is enriched further with detailed climate, material, and room usage data, to accurately reflect its performance under the influences of year-round occupancy and weather conditions. This complete NMT-BEM-model and its application for energy simulation are illustrated in Figure 4.

3.2 Energy Simulation for Thermal Comfort and Visual Comfort

Thermal comfort simulation data was evaluated using the *Predicted Mean Vote* (PMV) model [12], to determine the percentage of occupants likely to feel uncomfortable at any given time. While a detailed thermal comfort analysis was done, its elaboration would be beyond the scope of this paper, which focuses on daylight. One finding was, that excessive solar heat gains in summer, which elevate surface and operative temperatures are a primary cause of discomfort. Design modifications such as the addition of external shading devices may alleviate this situation, but they also influence natural daylight and visual comfort. An annual daylight analysis was conducted to determine the *Useful Daylight Illuminance* (UDI) for all rooms requiring daylight access. UDI was defined as the percentage of occupied hours within 300–3000 lux, per DIN EN 17037[13]. The *UDI low* and *UDI up* values are used to quantify under- and over-illumination (glare).

First, UDI results were averaged by key room types. Table 1 shows this analysis for treatment rooms (T.), reception areas (R.), office spaces (O.) and patient waiting areas (W.), as well as average UDI values for all daylit rooms in the entire building (B.). However, this averaging can mask significant differences among individual rooms of the same type, due to variations in room geometry, façade orientation and location. For reception areas, the overall UDI values were poor. While some were

Table 1: Reference Design

	UDI	UDI up	UDI low
T.	85.0	9.8	5.2
R.	36.1	1.7	62.2
O.	80.2	13.8	6.0
W.	83.5	11.0	5.5
B.	66.0	11.3	22.7

Table 2: Vertical Louvre Shades

	UDI	UDI up	UDI low
T.	87.2 (+2.2)	7.1 (-2.7)	5.7 (+0.5)
R.	36.1	1.7	62.2
O.	83.7 (+3.5)	8.9 (-4.9)	7.4 (+1.4)
W.	87.6 (+4.1)	5.3 (-5.7)	7.1 (+1.6)
B.	68.4 (+2.4)	8.0 (-3.3)	23.6 (+0.9)

Table 3: Design Change

	UDI	UDI up	UDI low
T.	83.6 (-1.4)	10.0 (+0.2)	6.4 (+1.2)
R.	59.0 (+22.9)	3.8 (+2.1)	37.2 (-25.0)
O.	76.9 (-3.3)	17.2 (+3.4)	6.0
W.	79.9 (-3.6)	15.2 (+4.2)	4.9 (-0.6)
B.	70.9 (+4.9)	12.0 (+0.7)	17.1 (-5.6)

Table 4: Design Change with Vertical Louvre Shades

	UDI	UDI up	UDI low
T.	85.5 (+0.5)	7.2 (-2.6)	7.3 (+2.1)
R.	58.9 (+22.8)	3.0 (+1.3)	38.1 (-24.1)
O.	79.6 (-0.6)	13.3 (-0.5)	7.1 (+1.1)
W.	84.8 (+1.3)	9.2 (-1.8)	6.0 (+0.5)
B.	73.2 (+7.2)	8.6 (-2.7)	18.2 (-4.5)

benefiting from a favourable northeastern orientation providing lots of daylight without glare, others received no daylight at all.

To address these discrepancies visual comfort was analysed in more detail, by employing UDI heatmaps that reveal localised variations in daylight distribution. This granular analysis was used to inform targeted interventions, with the goal of improving daylight provision to the unlit reception areas, as well as reduce glare and solar heating in the rooms along the southern façades.

3.3 Implementation of Design Changes

Based on the analyses, external shading and spatial reconfiguration were tested as design measures to improve daylight and comfort. Several different shading configurations on the southeast and southwest façades were simulated, including overhangs and horizontal and vertical louvres. North-angled vertical louvres performed best, reducing glare significantly, while still maintaining daylight availability, as can be seen in the significant reduction of UDI up values with only minor increases of UDI low (see Table 2).

In a separate simulation study, the floor plan was revised by recessing an exterior wall to create an open area with daylight access in a hallway and relocating the reception areas into this space, requiring other rooms to be shifted to make space for this change. The revision improved daylight quality in the reception areas by a much larger margin than the addition of sunshades did for the other room types. However, it also increased glare elsewhere (Table 3). Only by combining the floor plan revision with the addition of louvre shades (Table 4), could the requirements of both too-dark and too bright areas be balanced out, resulting in overall UDI improvements across most room types.

3.4 Room Dashboards

Analyses of thermal comfort and daylight showed that performance must often be evaluated at the level of specific rooms. To maximise the value of NMT–BEM in design practice, the interpretation of such granular simulation results therefore has to be simplified. Surveys indicate that designers value

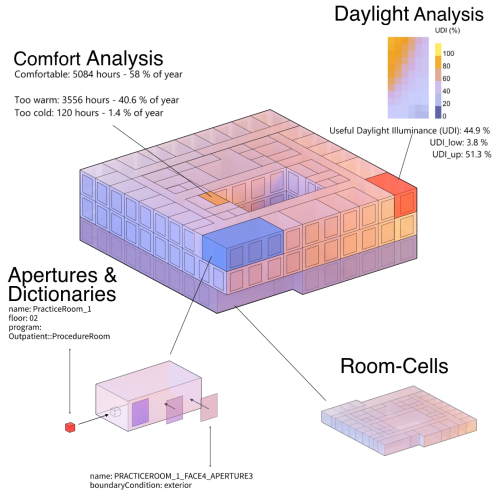


Figure 4: NMT-model with semantically enriched room Cells and BEM data

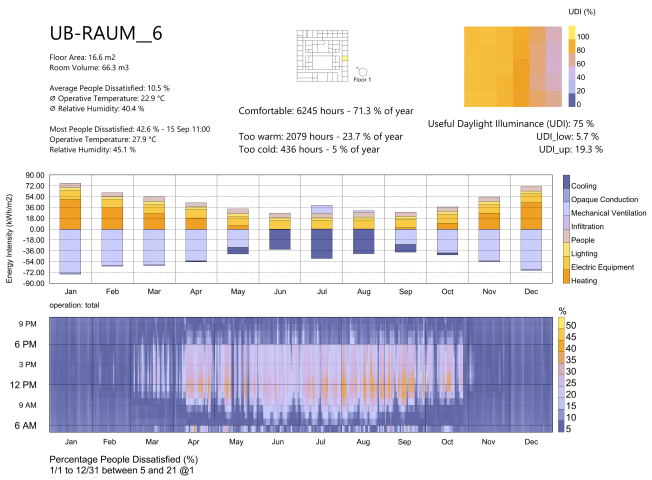


Figure 5: Exemplary dashboard of an examination room

tools which present study data in a visually accessible manner and frequently note a lack of automated result visualisation [14]. Result dashboards have been proposed as a solution to fill this gap [15].

To address this, a prototype dashboard was implemented focusing on thermal comfort and daylight metrics. The tool aggregates all simulation outputs, filters them by room, and presents key metrics with clear charts and concise text (see Figure 5). This approach enables rapid identification of underperforming spaces by translating complex results into intuitive, visual summaries, making evidence-based design changes more accessible and easier to validate.

4 Discussion

The study findings indicate that, while both measures are impactful, early-stage changes to the building volume and floor plan have a stronger influence on daylight quality than later-stage refinements to the shading concept. The results therefore demonstrate that the proposed NMT–BEM workflow achieved its intended outcome: lightweight room-cell models were generated rapidly and used to test multiple design alternatives within the conceptual phase. The use of NMT facilitated quick reconfiguration of the model, avoiding the time-consuming remodelling often associated with conventional BIM-to-BEM workflows.

Several advantages of the workflow can be identified. NMT ensured geometrically clean, watertight room models, eliminating common interoperability errors in BIM-based energy models. The semantic enrichment of room-cells allowed for detailed simulation, taking into account the function specific influences and requirements for indoor comfort and making targeted evaluation possible. The prototype dashboard further illustrated how room-level performance metrics can be communicated in a clear and accessible format, addressing known gaps in the usability of building performance tools.

At the same time, limitations must be acknowledged. The workflow relies on a highly-dependent, "proof-of-concept" software stack, which combines an in-development framework (Topologic) with open-source tools (Blender) and multiple plug-ins for commercial authoring software (Grasshopper, Ladybug/Honeybee). This may pose a barrier for practitioners due to complexity and lack of documen-

tation. The robust NMT modelling strategy developed here leveraged the regularity and intermediate size of the reference design; for larger or more geometrically complex buildings, additional methods may be needed to maintain stability.

Overall, the discussion highlights that the integration of NMT and BEM can provide timely, room-specific performance feedback that supports evidence-based decisions in conceptual design, but also that further refinement and broader testing are required for the workflow to be adopted in professional practice.

5 Conclusion

An integrated workflow was established that used non-manifold topology to generate lightweight room-cell models and connect them with established energy and daylight simulation tools. The case study demonstrated that such models enable targeted design refinements during early phases, for example by improving daylight distribution through spatial reconfiguration and shading. By focusing on essential spatial properties, NMT provided simulation-ready models that bridged the gap between conceptual design exploration and performance evaluation without the need for object-based BIM models. The introduction of room dashboards provided a robust interface for the extraction of detailed, room-specific performance data, effectively bridging the gap between complex simulation outputs and practical design decision making.

The findings highlight the potential of NMT to improve interoperability in the planning process and complement conventional BIM by facilitating early, evidence-based decisions while retaining compatibility with later, detailed design stages. Future work should focus on streamlining the workflow, expanding it to complex geometries and building types, and further integrating thermal comfort alongside daylighting to provide a more comprehensive evaluation. Beyond technical advances, the value of this contribution lies in enabling architects and planners to make better-informed decisions when they matter most. By providing timely design feedback in the early design stage, the approach facilitates healthier, more sustainable, and more energy-efficient buildings, benefits that extend directly to building occupants, clients, and communities.

Acknowledgements

I want to thank the office of *gmp Architekten von Gerkan, Marg und Partner*, Aachen for the permission to use their competition design for the "Medical Care Centre" in Achern, Baden-Württemberg as a case study and for kindly providing plans and documentation about the project.

References

[1] T. El-Khouly and O. Abdelhalim, “Preserving conceptual design integrity: Strategies for enhancing interoperability in architectural digital design workflows”, en, *Scientific Reports*, vol. 14, no. 1, Dec. 2024. DOI: 10.1038/s41598-024-78640-8

[2] C. Eastman, “Automated Assessment of Early Concept Designs”, *Architectural Design*, vol. 79, no. 2, 2009. DOI: 10.1002/ad.851

[3] P. Sanguinetti, S. Abdelmohsen, J. Lee, J. Lee, H. Sheward, and C. Eastman, “General system architecture for BIM: An integrated approach for design and analysis”, *Advanced Engineering Informatics*, vol. 26, no. 2, Apr. 2012. DOI: 10.1016/j.aei.2011.12.001

[4] J. W. Ouellette, P. Woodard, and M. Bekboliev, “Information Delivery Manual (IDM) Development for Building Information Modelling (BIM) and Building Energy Modelling (BEM) Workflows”, buildingSmart International ltd., Tech. Rep., Aug. 2022.

[5] R. Aish and A. Pratap, “Spatial Information Modeling of Buildings Using Non-Manifold Topology with ASM and DesignScript”, in *Advances in Architectural Geometry 2012*, L. Hesselgren, S. Sharma, J. Wallner, N. Baldassini, P. Bompas, and J. Raynaud, editors, Vienna: Springer Vienna, 2013. DOI: 10.1007/978-3-7091-1251-9_1

[6] S.-U. Lee, M.-I. Roh, J.-H. Cha, and K.-Y. Lee, “Ship compartment modeling based on a non-manifold polyhedron modeling kernel”, *Advances in Engineering Software*, vol. 40, no. 5, May 2009. DOI: 10.1016/j.advengsoft.2007.12.001

[7] M. Kiumarse Zamanian, S. J. Fenves, and E. Levent Gursoz, “Representing spatial abstractions of constructed facilities”, *Building and Environment*, vol. 27, no. 2, Apr. 1992. DOI: 10.1016/0360-1323(92)90024-J

[8] N. Wardhana, A. Chatzivasileiadi, W. Jabi, R. Aish, and S. Lannon, “Bespoke geometric glazing design for building energy performance analysis”, *FME Transactions*, vol. 47, no. 2, 2019. DOI: 10.5937/fmet1902370W

[9] A. Chatzivasileiadi, W. Jabi, S. Lannon, and N. M. Wardhana, “Addressing Pathways to Energy Modelling Through Non-Manifold Topology”, in *Proceedings of the 2018 Symposium on Simulation for Architecture and Urban Design (SimAUD 2018)*, Delft, Netherlands: Society for Modeling and Simulation International (SCS), 2018. DOI: 10.22360/SimAUD.2018.SimAUD.005

[10] A. Chatzivasileiadi, N. Wardhana, W. Jabi, R. Aish, and S. Lannon, “A Review of 3D Solid Modeling Software Libraries for Non-Manifold Modeling”, in *Proceedings of CAD’18*, CAD Solutions LLC, Jul. 2018. DOI: 10.14733/cadconfP.2018.59-65

[11] W. Jabi, *Wassimj/topologicpy*, Jul. 2025. Accessed: Jul. 6, 2025. [Online]. Available: https://github.com/wassimj/topologicpy

[12] P. O. Fanger, “Thermal comfort. Analysis and applications in environmental engineering.”, *Thermal comfort. Analysis and applications in environmental engineering.*, 1970.

[13] DIN EN 17037, “Tageslicht in Gebäuden”, Tech. Rep., May 2022.

[14] S. Attia, M. Hamdy, W. O’Brien, and S. Carlucci, “Assessing gaps and needs for integrating building performance optimization tools in net zero energy buildings design”, *Energy and Buildings*, vol. 60, May 2013. DOI: 10.1016/j.enbuild.2013.01.016

[15] M. Gadelhak, W. Lang, and F. Petzold, “A Visualization Dashboard and Decision Support Tool for Building Integrated Performance Optimization”, Sep. 2017. DOI: 10.52842/conf.ecaade.2017.1.719

Ontology-Driven GPT-4 Question Answering on Construction Standards

Merna Emara[1]
[1]Institut für Bauinformatik, Technische Universität Dresden, Dresden, Germany
E-mail(s): merna.emara@tu-dresden.de

Abstract: The German construction industry is governed by a complex and fragmented set of standards and technical regulations, often resulting in delays, high costs, and administrative overhead due to overregulation. This research proposes a framework that leverages semantic technologies and large language models (LLMs) to improve access to interpretation of construction standards. The approach begins with manual analysis of regulatory documents, which are structured into CSV files and transformed into RDF triples using the R2RML mapping language. These triples are integrated into a domain-specific ontology that captures key concepts such as building types, project phases, and material systems. To make compliance information more accessible, large language models (LLMs) such as GPT-4 are employed to interpret natural language queries, reason over structured ontology segments, and generate accurate responses without requiring formal SPARQL syntax. This approach combines semantics with the intuitive power of generative AI. Initial results show successful transformation into 1844 RDF triples, with LLM-generated answers evaluated against SPARQL-based ground truth.

Keywords: Semantic Web, Large Language Models (LLMs), Construction Standards

DOI: 10.18154/RWTH-CONV-254874. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

The construction industry relies on a complex collection of codes and standards that specify technical requirements and specifications to ensure aspects such as structural safety, accessibility, and interoperability in building's lifecycle. In Germany, the Building law term is largely divided into public and private building law such as the DIN (Deutsches Institut für Normung) and VDI (Verein Deutscher Ingenieure) standards. These texts are often written in dense, domain-specific language, which create challenges for digital integration [1]. To ensure that a building design meets regulatory requirements is a repetitive task, largely carried out manually, and it consumes significant money and time [2]. As Overregulation and bureaucracy have a direct impact on the cost of construction projects [3] and the industry increasingly moves toward digitization and Building Information Modeling (BIM), the need to make these standards machine-interpretable has become critical.

One commonly known method for representing standards in a machine-readable format is through the use of ontologies. Semantic web technologies have become increasingly popular in the AEC

domain as it provides structured, formal representations of domain knowledge and facilitate data interoperability, automated reasoning, and support intelligent querying [4]. Projects such as IfcOWL [5] and BOT (Building Topology Ontology) [6] have demonstrated how domain-specific ontologies can support semantic integration in BIM workflows. However, the systematic transformation of national regulatory standards into ontologies remains a largely open challenge due to the heterogeneity and complexity of the source documents.

To address the challenges of the initial transformation of textual or tabular data into semantic representations, the W3C's R2RML (RDB to RDF Mapping Language) specification offers a declarative approach for mapping structured data from relational databases to RDF datasets [7]. This language allows for flexible customization of mappings, enabling users to define the transformation of relational data into RDF according to their needs [8]. Simultaneously, the emergence of large language models (LLMs) such as GPT-4 [9] has created new opportunities to interact with semantic data through natural language. LLMs can assist users in querying, interpreting, and navigating ontologies without requiring deep technical expertise. Integrating LLMs into semantic systems can bridge the gap between human users and complex data models, especially for users without expertise in formal query languages like SPARQL. Recent studies have investigated the use of LLMs for tasks such as SPARQL generation, RDF triple completion, and question answering over knowledge graphs [10] . However, the use of LLMs for reasoning over ontologies derived from regulatory in the AEC sector has not been extensively explored or evaluated. This research develops a pipeline that converts German construction standards into ontologies and uses GPT-4 to answer questions over them with quantitative evaluation of the results. It aims to improve digital access to regulatory knowledge and support automation in construction.

2 Related studies

The construction industry stakeholders deal with a large number of regulations. Over time, several products and services have emerged to manage relevant documentation. Three common platforms include Nautos, a standards management system by Beuth Verlag, offering access to 2.6 million records from around 30 countries countries (e.g., Deutsches Institut für Normung DIN, Verein Deutscher Ingenieure VDI) [11], but its filtering criteria is restricted to basic metadata like issue date or publisher. Similarly, the VOB platform organizes regulations by activity, object type, or material but lacks comprehensive search functionalities. Lastly, the Building Standards Lexicon allows users to search across 20,000 sections from over 2,000 standards and guidelines [12]. These platforms mainly register standards without offering comprehensive, automated filtering, leaving users responsible for manually identifying relevant documents for their projects.

As digital transformation continues to reshape the Architecture, Engineering, and Construction (AEC) sector, there is an increasing need to convert such standards into machine-readable formats that can be integrated into software systems, particularly in Building Information Modeling (BIM) environments. Several research has explored the digitization of construction standards into ontologies to improve interoperability, compliance checking, and information retrieval. Efforts such as IfcOWL [5] and the Building Topology Ontology (BOT) [13] have demonstrated how building models and spatial relationships can be semantically represented for querying and reasoning. Zhou et al. [14] presented a first-order logic-

based information representation schema that enables automated regulatory compliance checking of building designs. Their research demonstrated the effectiveness of ontology based approaches for regulatory reasoning, achieving high recall and precision in detecting noncompliance instances. Buttgereit et al. [15] proposed a prototype for automated filtering of project-specific technical building regulations, leveraging IDS and IFC-compliant models to link building components with relevant rules for efficient compliance verification.

The emergence of Artificial Intelligence (AI), particularly through Large Language Models (LLMs), represents a breakthrough in automating the complex task of interpreting and structuring regulatory standards in the construction domain. Advanced models like LinkBERT [10] further improve document-level reasoning by capturing link semantics. In parallel, the rise of large language models (LLMs) such as GPT-4 [9] has created powerful tools for interacting with complex data using natural language. LLMs have been shown to perform well in tasks like SPARQL query generation, knowledge graph question answering, and ontology navigation [16]. By incorporating these models into semantic pipelines, it becomes possible to query domain-specific ontologies derived from regulatory content without requiring technical users to understand RDF or SPARQL. Allemang [17] presents a dual-component approach to enhance the accuracy of LLM-powered question-answering systems through ontology checks and repairs. The findings emphasize the importance of investing in metadata and ontologies for improved accuracy in real-world applications.

3 Methodology

This paper introduces a methodology for information retrieval by integrating ontological representations of technical standards with large language models (LLMs). The main objective is to semantically structure technical regulations to enable intelligent filtering based on project-specific requirements, and investigate the potential of LLMs for natural language questions answering over the knowledge base.

The methodology begins with a manual analysis of regulatory documents, most commonly published in PDF, XML, or HTML formats. While XML offers structured data benefits, among 16 federal states, only Bavaria and Berlin currently provide building regulations in XML format. As a result, the set of rules is manually extracted to capture metadata and convert the application area into a suitable format. Key regulatory information is extracted and tabulated, focusing on capturing metadata and defining the scope of application for each technical rule. Subsequently, this structured data is semantically mapped to an RDF ontology using the R2RML standard. This mapping defines relationships among rules, their scopes, and applicability criteria such as service types, system types, building materials, and document dependencies. The resulting ontology forms a machine interpretable model of the regulatory content. Then the scope is extended to employing GPT-4 in answering natural language questions about the ontology. Finally, the quality and accuracy of LLM-generated responses are quantitatively evaluated through comparison with a gold standard reference.

4 implementation

4.1 Data preparation

The scope of implementation is focused on selected parts of the MVV TB specifically Part A1 (Structural Stability) and Part A2 (Fire Protection). The Model Administrative Regulation for Technical Building Regulations (Muster-Verwaltungsvorschrift Technische Baubestimmungen) (MVV TB) is published by the German Institute for Building Technology (DIBt) in coordination with the federal state building authorities. Each section typically includes introductory remarks followed by a tabular list of building regulations. These entries consist of a unique ID, a textual scope of application, and abbreviated references to applicable technical standards. For this study, all applicable rules from Part A1 and A2 were analyzed. The analysis focused on the "Anwendungsbereich" (application scope) or "Geltungsbereich" (area of validity). When explicit sections were missing, the scope was inferred from section titles. The manual analysis identified four recurring classification criteria: (1) service description (Leistungsbeschreibung), specifying the type of activity governed by the rule; (2) system type (Anlagentyp), indicating the applicable phase; (3) building materials (Verwendete Baumaterialien), listing the relevant materials regulated by the; and (4) document dependency (Dokumentabhängigkeit), denoting cross-references to other standards. The analysis of the regulation is carried out manually and the extracted content is formalized into a tabular structure, identifying key regulatory criteria to support machine interpretation.

4.2 Knowledge Base Construction: R2RML Mapping

To transform the tabulated data into a semantic representation, the R2RML (RDB to RDF Mapping Language) standard was adopted. R2RML enables the mapping of relational or tabular data to RDF by specifying triples composed of subjects, predicates, and objects. Recommended by the W3C, R2RML provides a standardized approach for expressing how relational data should be mapped to RDF datasets [8].

The study utilizes R2RML-F, an extension of R2RML tailored for mapping CSV files into RDF format. This tool, developed in Java and interacts with CSV files using H2 Database Engine, reads CSV files and generates RDF outputs in Turtle format. On the conversion process, column headers have to be transformed to uppercase, and several files should be a available for execution: the input CSV file, a mapping file (.ttl), a configuration file, the output RDF file, and the associated JAR executable [8]. This mapping process facilitates the integration of the analyised technical rules into an ontology. Each row in the CSV was treated as an instance of the TechnischeRegel class, and mapping files were defined to capture relationships among properties. The output was stored in Turtle format and merged with a predefined ontology using consistent IRI naming. After successful transformation and merging, the RDF data is incorporated into the knowledge base, enabling structured retrieval and visualization through SPARQL.

Our knowledge base represents standards as instances of the class reg:TechnicalRule. Each rule is linked to multiple domain facets that capture its scope and applicability. For example, the property reg:hasObjectType associates a rule with the relevant building or element types (such as Buildings, Civil engineering structures, or Decke), while reg:hasMaterial specifies the materials covered (e.g., Beton, Betonstahl, Mörtel). The property reg:hasProjectPhase connects rules to the HOAI project

phases (LP1–LP9), and reg:hasServiceProfile links them to the applicable professional discipline, such as Tragwerksplanung (structural design). In addition, inter-rule dependencies are captured through reg:dependsOn and reg:dependencyType, which denote relationships like Berichtigung (correction) or National Addition. As a concrete example, the rule https://w3id.org/regelwerk/tr#DIN_1045-2:2008-08 (“DIN 1045-2:2008-08”) is modeled as a reg:TechnicalRule associated with the object types Buildings and Civil engineering structures, the material Beton, the project phases LP1 to LP5, and the service profile Structural design. This information is encoded in RDF as follows:

Algorithm 1: RDF/Turtle representation of the technical rule DIN 1045-2:2008-08.

```
prefix reg:  <https://w3id.org/regelwerk#> .
@prefix tr:   <https://w3id.org/regelwerk/tr#> .
@prefix obj:  <https://w3id.org/regelwerk/object#> .
@prefix mat:  <https://w3id.org/regelwerk/material#> .
@prefix prj:  <https://w3id.org/regelwerk/project#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .

tr:DIN_1045-2:2008-08 a reg:TechnicalRule ;
  rdfs:label "DIN 1045-2:2008-08" ;
  reg:hasObjectType obj:Gebaeude , obj:Ingenieurbauwerk ;
  reg:hasMaterial   mat:Beton ;
  reg:hasProjectPhase prj:LP1 , prj:LP2 , prj:LP3 , prj:LP4 , prj:LP5 ;
  reg:hasServiceProfile prj:Tragwerksplanung .
```

4.3 LLM-Based Querying

To bridge the gap between ontologies and humane interaction, GPT-4 was employed to interpret natural language queries and retrieve relevant knowledge from the ontology. Instead of translating questions into SPARQL, this method leverages GPT-4’s reasoning capabilities to infer structured answers directly from the ontology.

After mapping German construction standards to RDF triples using R2RML and storing them in Turtle format, the ontology was loaded into an RDF graph. Each resource (subject node) was converted into a human-readable entity card that captures its resource’s identifiers (IRI), labels, types, and a limited set of descriptive triples (facts). These entity cards serve as natural language summaries of ontology entities. To enable efficient retrieval, all entity cards were embedded using OpenAI’s text-embedding-3-large model and indexed with FAISS (Facebook AI Similarity Search). This allows user questions to be embedded in the same vector space, and the most semantically relevant entity cards can be retrieved using similarity search. By narrowing the ontology context to only the top-k relevant cards, the system ensures scalability and reduces noise in the LLM’s input.

The retrieved entity cards, along with the user’s question, are then passed to GPT-4 using a dedicated prompt. Algorithm 2 illustrates a snippet of our ontology QA assistant implementation. The code defines the system-level instructions, submits the entity cards and question to the LLM, and parses the response into structured JSON containing two fields: the predicted IRIs and the supporting RDF triples.

Algorithm 2: Ontology QA assistant implementation in Python

```
ANSWER_SYSTEM = """You are a careful ontology QA assistant.
Use ONLY the provided entity cards.
Return JSON with:
- predicted_iris: array of IRIs (strings).
- supporting_triples: array of [s,p,o] (strings) taken from the cards.
"""
question = "What technical standards and codes apply to buildings made of
    reinforced concrete, across all project phases?"
cards_k = retrieve_cards(question, k=15)
pred_json = ask_llm(question, cards_k)
pred_iris = pred_json.get("predicted_iris", [])
supporting = pred_json.get("supporting_triples", [])
```

The model is instructed to act as an ontology QA assistant and to return structured JSON output containing:

- predicted_iris: a list of IRIs relevant to the question
- supporting_triples: a set of RDF triples (subject, predicate, object) extracted from the provided entity cards that justify the prediction

For example, when asked “What technical standards and codes apply to geotechnical engineering where the materials include shotcrete (Spritzbeton)?”, the system retrieves entity cards describing reinforced concrete and related standards, and GPT-4 produces a set of IRIs together with supporting triples explaining why they were selected.

5 Results

The proposed pipeline successfully transformed German construction standards from CSV files into structured RDF triples using R2RML mappings. This process generated a total of 1,844 triples, forming the ontology used for querying and evaluation.

To evaluate the effectiveness of the LLM-based querying mechanism, outputs generated by GPT-4 were compared against ground truth answers obtained through direct SPARQL queries in protege over the same ontology. This comparative analysis aimed to measure the accuracy and precision of the natural language interface in retrieving the correct IRIs for user queries.

Preliminary evaluation on a set of test queries yielded a precision of 0.46, recall of 0.49, and F1 score of 0.47. While these results indicate that the model can retrieve relevant standards, they also highlight challenges in achieving exact matches.

For example, in response to the query: “What technical standards and codes apply to geotechnical engineering (Geotechnik) for bored piles (Bohrpfahl)?” GPT-4 produced the following results:

- DIN_EN_1536:2010-12 (correct match for bored piles),
- DIN_EN_1993-5:2010-12 (retrieved because its description includes the substring “pfahl”),
- DIN_EN_12699:2001-05 (retrieved because its description includes “Verdrängungspfahl”)

These cases illustrate a recurring error: the system sometimes retrieves partial string matches (e.g., any entity containing “pfahl”) rather than restricting results to exact semantic matches.

To address these limitations, two potential enhancements were identified. First, the retrieval logic can be refined to enforce exact matches rather than allowing partial substring overlaps, thereby reducing false positives. Second, the effectiveness of the system can be improved through more precise query formulation. For example, instead of asking:

“What technical standards and codes apply to the structural design of industrial buildings, regardless of construction material?”

a more specific formulation such as:

“Which technical rules (TechnicalRule) apply to factory/industrial buildings—i.e., rules whose object-type label includes ‘Industriebau’, ‘Fabrik’, or ‘industrial’?”

would likely yield more accurate results. However, this approach assumes that end users are familiar with the ontology’s underlying classes and object labels, which is often unrealistic in practice.

6 Conclusion and future work

This study introduces a framework that combines semantic technologies and large language models (LLMs) to help reason over structured construction data. The methodology involved manually analyzing and structuring regulatory documents (MVV TB Parts A1 and A2) into tabular CSV format, which was then transformed into RDF triples using R2RML and integrated into a domain-specific ontology. The resulting knowledge base supports reasoning over construction regulations through large language models (LLMs) such as GPT-4, enabling users to submit project-specific queries without the need to formulate SPARQL statements. GPT-4 interprets the queries and provides informed responses based on the technical standards.

Despite these promising outcomes, several limitations remain. The workflow currently depends on manual interpretation and structuring of regulatory texts, which introduces potential inconsistencies and reduces scalability. Future work will focus on automating or semi-automating this process through advanced natural language processing (NLP) techniques, including named entity recognition, relation extraction, and legal text classification to identify scope statements and inter-rule dependencies within building codes. Moreover, while GPT-4 demonstrates the ability to reason over segments of the ontology, its answers remain probabilistic and may occasionally diverge from strict rule-based logic.

References

[1] E. P. Karan, J. Irizarry, and J. Haymaker, “Bim and gis integration and interoperability based on semantic web technology”, *J. Comput. Civ. Eng.*, vol. 30, no. 3, p. 04 015 043, 2016. DOI: 10.1061/(ASCE)CP.1943-5487.0000519

[2] A. Borrmann, M. König, C. Koch, and J. Beetz, *Building Information Modeling*. Cham: Springer International Publishing, 2018.

[3] D. Walberg, T. Gniechwitz, K. Paare, and Schulze, *Status und prognose: So baut deutschland - so wohnt deutschland. der chancen-check für den wohnungsbau*, 2023.

[4] J. Zhang and N. M. El-Gohary, “Semantic-based logic representation and reasoning for automated regulatory compliance checking”, *J. Comput. Civ. Eng.*, vol. 31, no. 1, p. 04 016 037, 2017. DOI: 10.1061/(ASCE)CP.1943-5487.0000583

[5] P. Pauwels, D. Van Deursen, R. Verstraeten, J. De Roo, et al., *Ifcowl: A case of transforming express schemas into ontologies*, https://github.com/openBIMstandards/ifcOWL?utm_source=chatgpt.com, 2011.

[6] M. H. Rasmussen, M. Lefrançois, G. F. Schneider, and P. Pauwels, *Bot: The building topology ontology of the w3c linked building data group*, https://www.semantic-web-journal.net/content/bot-building-topology-ontology-w3c-linked-building-data-group-0, 2019.

[7] W3C, *R2rml: Rdb to rdf mapping language*, https://www.w3.org/TR/r2rml/, 2012.

[8] S. Das, S. Sundara, and R. Cyganiak, *R2rml: Rdb to rdf mapping language*, https://www.w3.org/TR/r2rml/, 2012.

[9] OpenAI, *Gpt-4 technical report*, https://openai.com/index/gpt-4-research/, 2023.

[10] M. Yasunaga, J. Leskovec, and P. Liang, *Linkbert: Pretraining language models with document links*, 2022.

[11] Beuth Verlag GmbH, *Nautos: Beuths normen-management-lösung - beuth.de*, https://www.beuth.de/de/normen-management/nautos, 2023.

[12] P. Flach, *Machine Learning*. Cambridge University Press, 2012.

[13] P. Pauwels and A. Roxin, “Simplebim: From full ifcowl graphs to simplified building graphs”, in *Ework and Ebusiness in Architecture, Engineering and Construction*, S. Christodoulou and R. Scherer, editors, CRC Press, 2016, pp. 11–18. [Online]. Available: http://hdl.handle.net/1854/LU-8041826

[14] Y. Zhou, J. She, Y. Huang, L. Li, L. Zhang, and J. Zhang, “A design for safety (dfs) semantic framework based on nlp for automated compliance checking using bim: The case of china”, *Buildings*, vol. 12, no. 6, p. 780, 2022.

[15] J. Stührenberg, T. Al-Zuriqat, and C. C. Geck, “Tagungsband des 35. forum bauinformatik 2024”, in *35. Forum Bauinformatik 2024*, 2024. DOI: 10.15480/882.13486

[16] S. Qiao et al., “Reasoning with language model prompting: A survey”, in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2023, pp. 5368–5393. DOI: 10.18653/v1/2023.acl-long.294 [Online]. Available: https://aclanthology.org/2023.acl-long.294

[17] D. Allemang and J. Sequeda, *Increasing the llm accuracy for question answering: Ontologies to the rescue!*, http://arxiv.org/pdf/2405.11706v1, 2024.

Persistence and linking of geometric-topological data models in BIM and GIS

Sophie Lipskaya[1], Robert Jürgensen[1] and Maximilian Sternal[1]
[1]Chair of Computing in Civil Engineering, Technische Universität Berlin, Germany
E-mail(s): s.lipskaya@campus.tu-berlin.de, r.juergensen@campus.tu-berlin.de

Abstract: The integration of Building Information Modeling (BIM) and Geographic Information Systems (GIS) requires robust geometric-topological consistency to avoid the costly extract-transform-load and repair processes that burden current workflows. This paper addresses three critical requirements for sustainable geometric-topological models: (a) consistent modeling tools that maintain topological integrity throughout the modeling process, (b) persistent data structures with stable identification for specific geometric and topological entities such as faces and solids, and (c) linking strategies that preserve geometric and semantic relationships without relying on error-prone data conversion pipelines. This approach is presented through the Partition Platform, a prototype modeling tool that begins modeling processes in partitioned rather than empty space, always ensuring a valid partition of the Euclidian space and maintaining full adjacency information throughout the workflow. The study demonstrates how modern JSON-based schemas, specifically IFC5 and CityJSON, represent progress toward geometric-topological persistence. Building on these foundations, it is shown how JSON-LD can extend beyond semantic linking to establish explicit connections between geometric elements, enabling direct logical and topological relationships between faces and spatial features across BIM and GIS domains. This approach eliminates the need for lossy geometric translations, instead enabling lightweight interoperability through persistent geometric identifiers. The resulting framework supports robust data exchange for digital twins, infrastructure modeling, and spatial analysis applications where geometric-topological integrity is essential for reliable computational results.

Keywords: Building Information Modeling (BIM), Geometric Modeling, Topological Relationships, Linked Data, Geographic Information Systems (GIS)

DOI: 10.18154/RWTH-CONV-254875. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Geometry is the foundation of digital representations of the built environment, yet most of the common Building Information Modeling (BIM) tools show geometric and topological inconsistencies. Most geometric representations are not based on explicit topological entities, such as faces, but on procedural or mesh-based geometry, making boundaries ambiguous and preventing the storage of topological information such as inner faces bounding rooms and outer faces defining the building envelope. This limits both geometric clarity and the possibility of embedding topology directly in the model. Partially

this comes from the object-based modeling paradigm used in current geometric modeling tools. Thus, adjacency and connectivity are not inherent, but must be reconstructed from raw geometry. Many models also show geometric inconsistencies such as overlaps or gaps, and there is typically no stable identification of geometry entities, making reliable linking and long-term data management difficult.

Transformation-based integration between BIM and Geographic Information Systems (GIS) further increases these issues. Geometric simplification, which represents the most common form of loss (see Figure 1), often occurs when parametric BIM geometry is converted into simpler primitives [1], while topological relationships embedded in the source format may be lost or require reconstruction. Level of Detail (LoD) mismatches create further problems, where reconstructing detailed BIM from GIS data often produces semantically poor models [2].

The GIS side, however, shows some advantages. Formats such as CityJSON already provide explicit boundary-based geometry, maintain consistent hierarchies and identification of every face, which allows elements to be linked easily, although they lack the semantic richness and parametric flexibility of BIM.

Improving the situation requires face-based persistence to capture both geometry and inner–outer face relationships, modeling techniques that ensure geometric and topological integrity, approaches that link rather than transform data, and formats that persistently store identifiers with geometry. Connecting BIM and GIS is essential because BIM offers building-scale semantics and object-oriented geometry in high detail, while GIS focuses on large-scale spatial organization and analysis. Together, they can support holistic, multi-scale digital twins that support applications such as city-wide energy analysis, infrastructure planning across multiple responsibilities, and climate resilience assessment, combining building details with their wider spatial context.

This paper addresses these challenges by analyzing geometric and topological persistence in BIM and GIS, focusing on modern JSON-based serialization formats such as CityJSON and IFC5. For the generation of consistent geometric-topological entities, it introduces the Partition Platform, a prototype for space partitioning–based modeling with robust topology via cellular decomposition and pointer-based navigation, providing persistent identifiers. JSON-LD is used to demonstrate cross-domain linking strategies in two examples that retain original data structures while enabling modular integration. The paper concludes with an outlook on the use of linked data with robust geometry persistence in the future.

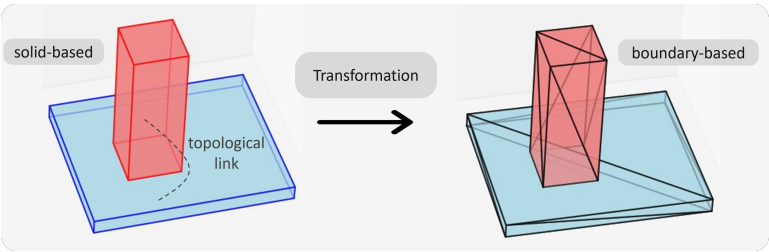


Figure 1: Information loss due to transformation from topological solid- to inconsistent mesh-based geometry.

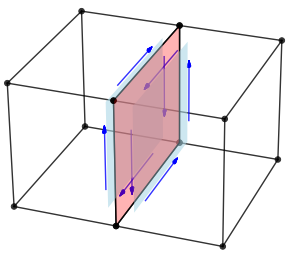


Figure 2: Geometric and topologigal model of common face (red) of two cubes in the Partition Platform.

2 Background

Understanding the geometric and topological persistence challenges in BIM-GIS integration requires examining the evolution of data models and serialization formats, the various approaches to format transformation, and emerging linked data alternatives. This background provides the foundation for understanding why current transformation approaches often result in information loss and how modern JSON-based formats, combined with robust topological modeling and linked data strategies, can address these limitations.

2.1 Introducing the Partition Platform

The Partition Platform (PP) is based on the partition of the entire unbounded Euclidean space from the beginning of the modeling process onwards. It defines clear polyhedral spatial boundaries to separate subspaces and maintain adjacency information and geometric robustness in every modeling step [3]. The topological pointer system successfully maintains edge-face and vertex-edge connectivity relationships, with cellular partitioning boundaries showing particular robustness in resolving modeling inconsistencies [4].

PP supports unbounded domains, as well as non-convex and multiply connected geometric configurations. It explicitly handles oriented and unoriented geometry for points, edges, faces, and volumes (cells). An example of a shared surface of two adjacent cubes in the Partition Platform is shown in Figure 2. The two cubes are a subset of the partition model. The face (red) shows the common geometry, and the blue facets are directed polygons with their normals pointing to the adjacent space (topology).

For persistent data storage, the topological model can be serialized into a nested JSON structure in which each topological element is uniquely identifiable by an ID. Coordinates are stored as a list of vertices, while topology and geometry are kept separate. Model reconstruction from this structure has a computational cost of order $O(n)$, with n as the total number of topological entities in the model.

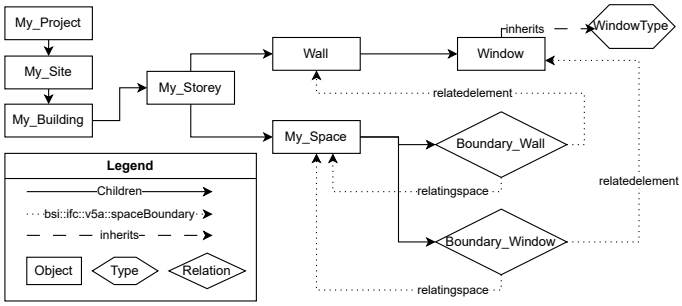


Figure 3: Hierarchical structure of IFC5 elements, illustrating parent-child and adjacency relationships

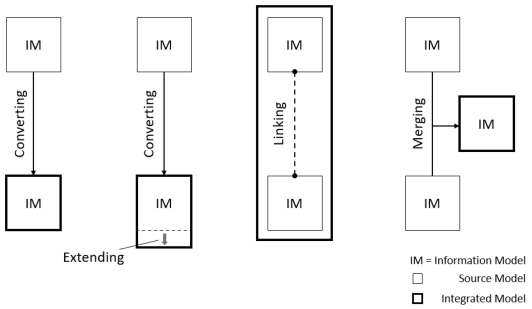


Figure 4: Different Information Model Strategies, with focus on linking [2]

2.2 Linked Data

Traditional approaches to BIM-GIS integration rely on format conversion to achieve interoperability, where data from one domain is transformed into data structures and geometric representations of another. Although conceptually straightforward, these transformation methodologies face, as mentioned above, significant challenges in preserving geometric accuracy and topological relationships throughout the conversion process.

The fundamental distinction between linking and conversion approaches lies in their treatment of source data and integration methodology (see Figure 4). Traditional conversion approaches create derivative models that represent information from one domain using the data structures of another, inevitably resulting in loss of information and structural adaptation [1]. In contrast, linking approaches maintain the original data models unchanged while establishing explicit cross-domain connections through definitions of external relationships [2], [5].

This way, the integrity of the source model data is preserved, which has several advantages for geometric and topological persistence. First, parametric geometric representations remain intact within their native formats, allowing continued model evolution and design iteration without degradation. Second, domain-specific topological relationships are preserved within their original structural contexts, maintaining the spatial hierarchies and connectivity information essential for domain-specific analysis. Third, the linking approach supports bidirectional integration without the asymmetric information loss characteristic of transformation workflows.

```
{  "path": "c8ecbf4c-e37a-4489-9133-15163
    b8a904e",
  "attributes":
  {
    "bsi::ifc::v5a::spaceBoundary": {
      "relatedelement": {
        "ref": "93791d5d-5beb-437b-b8ec-2
            f1f0ba4bf3b"
      },
      "relatingspace": {
        "ref": "e3035b71-bd9f-4cdc-86fd-
            b56e2f4605b6"
      }}}}
```

```
{  "@context": {
    "ifcx": "https://example.org/ifcx#",
    "cj": "https://example.org/cityjson#",
    "owl": "http://www.w3.org/2002/07/owl#"
  },
  "@graph": [
    { "@id": "ifcx:geometry424",
      "@type": "ifcx:WindowType",
      "owl:sameAs": "cj:window6"
    },
    { "@id": "cj:window6",
      "@type": "cj:Window",
      "owl:sameAs": "ifcx:geometry424" }]}
```

Figure 5: Side-by-side JSON representations: (left) topological relation, (right) linked window geometry in JSON-LD.

JSON-LD Implementation Requirements: JSON-LD employs triple-based structures specifying subject-predicate-object connections between elements across data sources, enabling definitions of spatial adjacency, geometric overlap, scale-based correspondence between BIM elements and GIS features (Figure 5, right), and explicit topological hierarchies that span domain boundaries, such as linking IFC spaces to CityJSON building rooms, as illustrated in Figure 7 with the green arrow. Its distributed nature allows data sources to remain in their original locations while maintaining logical integration through web-accessible relationship definitions [5]. Successful implementation depends on consistent topological models that preserve geometric precision and spatial relationship integrity across the linking framework, motivating specialized topological platforms for stable geometric foundations [3].

Implementation requires three fundamental components: unique identifiers for linkable elements, structured relationship definitions, and accessible data locations.

2.3 CityJSON and IFC5

BIM and GIS have traditionally relied on distinct data models and serialization approaches that reflect their different spatial scales and application domains. Although these formats have served their

respective communities effectively, their structural differences and geometric representation methods present significant challenges for cross-domain integration.

Established Formats: CityGML employs an XML-based serialization format designed for representing 3D city models with semantic information across multiple levels of detail. The format structures geometric data using boundary representation (B-rep) approaches, where buildings and urban features are represented through their surface boundaries. However, the XML-based structure creates a hierarchical tree representation that requires deep navigation to access geometric primitives, making programmatic access and web-based applications less efficient [6].

IFC, with its latest version 4.3, represents an even more complex hierarchical approach to storing BIM data, typically serialized using the STEP format where each line represents a specific object with its attributes. The IFC schema supports numerous geometric primitives including Constructive Solid Geometry (CSG), boundary representations, extrusions, swept surfaces, and parametric shapes, providing extensive flexibility for detailed building component representation [7]. However, this geometric richness comes at the cost of accessibility. The hierarchical referencing system, while powerful, creates barriers to efficient geometric extraction and spatial relationship analysis.

Modern JSON-Based Approaches: The limitations of XML and STEP-based formats have driven the development of JSON-based alternatives that prioritize structural simplicity and web compatibility while maintaining semantic richness. CityJSON version 2.0, standardized in 2023 [6], provides a simplified JSON representation of the CityGML data model. The format's flat structure allows direct access to geometric objects without deep tree navigation, making it significantly more accessible for web technologies and modern development environments.

IFC5, currently under alpha version development by buildingSMART, represents a complete reboot and fundamental shift toward modular, JSON-based BIM data representation [8]. Unlike previous attempts to create JSON encodings of IFC4 that merely translated the hierarchical schema to JSON syntax, IFC5 redesigns the data model itself to leverage JSON's strengths [9]. The new format employs a flat, modularized structure where object order within files becomes arbitrary since geometries are identified through unique IDs rather than local references. This modularity enables selective loading of specific data components based on application requirements [8]. Parent-child relationships are maintained through explicit reference systems rather than structural nesting [8]. JSON approaches support enhanced topological modeling through three distinct relationship types maintained via identifier-based references (see Figure 3). Topological relationships are explicitly maintained through space boundary objects linking spaces to elements via relatingSpace and relatedElement attributes, preserving interface geometry (see Figure 5 left).

IFC5's flat structure with unique object IDs enables direct geometric and semantic linking, while CityJSON's identifier system supports feature-level and face-based connections, contrasting favorably with IFC4's reference-based geometry system, which lacks explicit geometry identifiers necessary for external linking [10].

3 Examples

The following examples demonstrate practical applications of the concepts discussed in the background section, illustrating how robust topological modeling and linked data approaches can address common challenges in BIM-GIS integration. These examples highlight both the potential and limitations of current approaches while demonstrating the effectiveness of combining topological consistency with modular JSON-based linking strategies.

3.1 Topological Relationships and Geometric Consistency

Analysis of the buildingSMART `Hello-Wall` IFC5 example model [8] reveals three significant geometric inconsistencies that demonstrate the need for robust topological modeling and persistence. First, multiple vertices and triangles appear in the geometry mesh where the space body contains each point three times, resulting in redundant representation (24 vertices instead of 8 for a simple box). Second, rounding inaccuracies and unnecessary transformations create potential geometric clashes or gaps, such as the boundary wall positioned at $-2.3841858\mathrm{e}-8$ instead of the expected 0.0. Third, boundary surface mismatches occur as the `Boundary_Wall` consists of two triangles forming a rectangle, while the corresponding wall surface has several triangles due to window openings. The shape mismatch of the contact surfaces can be seen in the explosion graphic in Figure 6. The boundary surfaces are colored red and yellow for `Wall` and `Window` respectively.

By generating IFC5 objects from the Partition Platform or linking to corresponding topological model elements in the PP-JSON file via JSON-LD, consistency can be maintained while preserving the modular structure of the source data. Since not only solids (cells), but each topological entity (nodes, edges, faces) is identifiable in the Partition Platform, a correct adjacent surface with holes can be exported or linked to `Boundary_Wall` from the IFC5. This way, not only is the topological consistency guaranteed, but also the surfaces of neighboring elements would identically match.

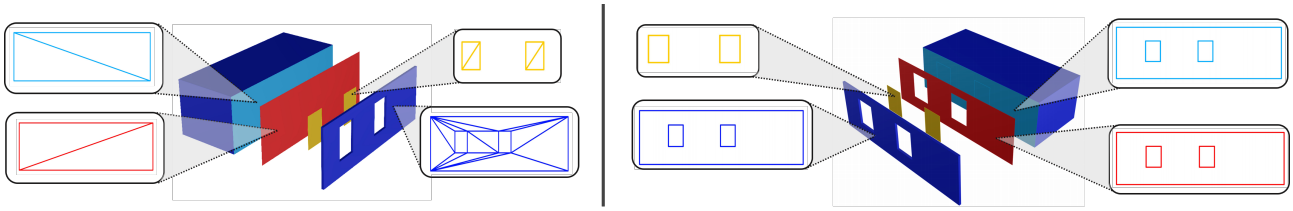


Figure 6: Explosion of selected objects of `Hello-Wall` IFC5 model [8] (left) and of respective objects in the Partition Platform (right)

3.2 BIM-GIS Linking with Modular JSON Structures

Using the IFC5 `Hello-Wall` model and a corresponding CityJSON representation created from [10], two linking approaches demonstrate the versatility of JSON-LD for BIM-GIS integration. The first approach links building elements at the object level, connecting the IFC space with the corresponding CityJSON `BuildingRoom` using a `owl:sameAs` relationship. This semantic linking maintains object identity across domains while preserving the original data structures in both files and is compatible with the current IFC standard, as object-level identifiers are already supported in IFC4. The second approach demonstrates geometry-level linking, connecting the detailed window geometry from the IFC5 file to the simplified window object in CityJSON, as shown in Figure 5, right. Unlike object-level

linking, this geometry-level linking is only possible with IFC5, since the current IFC4 standard does not assign unique identifiers to geometry resources. Thus, IFC5 enables applications that require detailed BIM geometry combined with GIS spatial context without a complete format transformation. The linking visualization in Figure 7 illustrates how the same building in different formats can be simultaneously accessed through domain-specific viewers while maintaining logical connections through JSON-LD relationships. In this figure, the green arrow represents the object-level link between the room entities in both domains, while the red arrow indicates the geometry-level link between the IFC5 window geometry and the corresponding CityJSON window object, as seen exemplarily on the right in Figure 5.

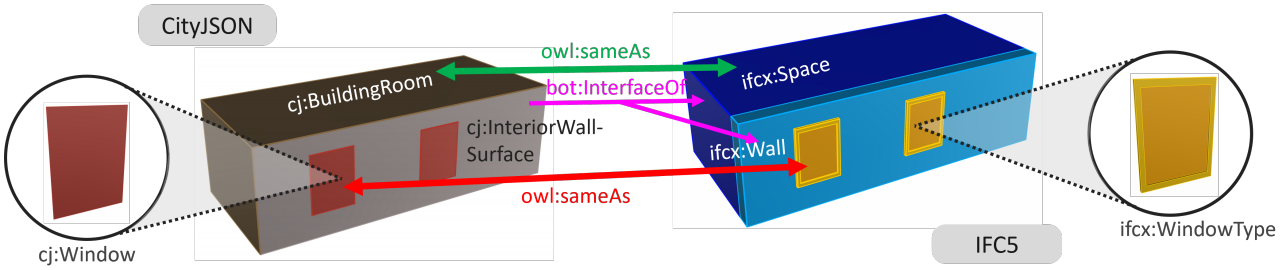


Figure 7: Visualization of linked elements: IFC viewer on the left and CityJSON viewer on the right

In the context of urban planning, another potential application of Linked Data is to establish direct connections between planned building models in BIM and the corresponding property features in GIS. This can be achieved by linking a property boundary and neighboring building connections, represented as a `WallSurface` in the CityJSON dataset, to the corresponding exterior wall geometry solid or boundary surface in an IFC5 model (blue and red meshes in Figure 6).

4 Conclusions and Outlook

This study demonstrates that modular JSON-based formats offer significant advantages for BIM-GIS integration over traditional hierarchical approaches. The transition to flat JSON structures with unique identifier systems, also for geometries, enables the possibility to add robust topological relationships and flexible cross-domain linking without information loss. The Partition Platform offers modeling geometry with topological consistency across serialization cycles while internally addressing geometric inconsistencies in current implementations that have also been shown in the minimal example presented in this paper. JSON-LD linking strategies preserve source model integrity while establishing semantic connections, representing a paradigm shift from transformation-based to preservation-based integration. This approach opens opportunities for integration with semantic web technologies and ISO 21597 (ICDD) standards [11]. The explicit linkable nature of geometries in IFC5 and CityJSON enables participation in broader knowledge graphs and standardized information containers, supporting digital twin and smart city applications.

Future research could focus on the integration of IndoorJSON for indoor navigation and connectivity modeling, extending the presented approach from city- and building-scale to interior environments. Another important aspect is the development and adoption of standardized linking vocabularies for spatial relationships, addressing data governance challenges in distributed linked data scenarios. The

Building Topology Ontology (BOT), developed by the W3C Building Data Group, already provides vocabulary for describing building topology [12]. The pink arrows in Figure (7) show such an application. The demonstrated approaches for persistent, identifiable geometric and topological data point toward building and urban models that maintain domain-specific strengths while participating in flexible, modular integration networks that are open to user-specific data models.

References

[1] C. Clemen, F. Gruner, and J. Pfeifer, “Infrastrukturdatenhaltung mit BIM und GIS”, Deutsches Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt, 2023. DOI: 10.48755/DZSF.230015.01

[2] S. F. Beck, “Context-sensitive linking of heterogeneous information models from the building and the urban domain”, Ph.D. dissertation, TU München, 2023.

[3] M. Sternal, “Polyhedral partition of the euclidean space for consistent spatial modeling of the built environment”, Ph.D. dissertation, TU Berlin, 2025. DOI: 10.14279/depositonce-23869

[4] W. Huhnt, M. Sternal, and P. J. Pahl, “Modeling bounded and unbounded space with polyhedra: Topology and operators for manifold cell complexes”, *Advanced Engineering Informatics*, vol. 54, Oct. 2022. DOI: 10.1016/j.aei.2022.101790

[5] P. Pauwels, A. Costin, and M. H. Rasmussen, “Knowledge graphs and linked data for the built environment”, in *Industry 4.0 for the Built Environment*, Cham: Springer International Publishing, 2022. DOI: 10.1007/978-3-030-82430-3_7

[6] OGC. “CityJSON community standard”, Accessed: Jul. 1, 2025. [Online]. Available: http://www.opengis.net/doc/CS/covjson/2.0

[7] F. Noardo, “Multisource spatial data integration for use cases applications”, *Transactions in GIS*, vol. 26, no. 7, pp. 2874–2913, Nov. 2022. DOI: 10.1111/tgis.12987

[8] buildingSMART International. “Ifc5 development – github repository”, Accessed: Jul. 1, 2025. [Online]. Available: https://github.com/buildingSMART/IFC5-development

[9] L. van Berlo et al., “Future of the industry foundation classes: Towards IFC 5”, *Proc. of the 38th International Conference of CIB W78*, pp. 123–137, 2021.

[10] H. İ. Şenol and T. Gökgöz, “Integration of building information modeling (BIM) and geographic information system (GIS): A new approach for IFC to CityJSON conversion”, *Earth Sci Inform*, vol. 17, Aug. 1, 2024. DOI: 10.1007/s12145-024-01343-1

[11] J. Krischler, P.-C. Schuler, J. Taraben, and C. Koch, “Using icdd for bim and gis integration in infrastructure”, *LDAC 2024: 12th Linked Data in Architecture and Construction Workshop, June 13–14, 2024, Bochum, Germany*, 2024.

[12] M. H. Rasmussen, M. Lefrançois, G. F. Schneider, and P. Pauwels, “BOT: The building topology ontology of the w3c linked building data group”, *Semantic Web*, vol. 12, pp. 143–161, 2020.

Image Captioning for Building Damage Using Deep Learning: Explaining Context in Overlapping Structural Defects

Vedant Girish Dalvi[1], Jakob Martin[1] and Timur Weilbach-Eyüboglu[1]
[1]Hochschule München University of Applied Sciences, Munich, Germany
E-mail(s): vedant_girish.dalvi@hm.edu, jakob.martin0@hm.edu, timur.weilbach-eyueboglu@hm.edu

Abstract: Different types of damage are found in building structures, which vary depending on the materials used, such as steel, concrete, and glass. Deep learning-based vision models have been widely employed in recent years to classify, detect, and segment these damages. However, damages often overlap in real-world scenarios, creating complex image segmentations. The intricate nature of construction damages is effectively described using captions generated by deep learning-based Vision-Language Models (VLMs), which provide insight into the impact of these damages on structural stability. This paper proposes an image-captioning web application for automating the documentation of construction damages, thereby increasing the efficiency of safety inspections, workflows, and maintenance. The application utilizes a fine-tuned Bootstrapping Language-Image Pre-training (BLIP) architecture to generate captions that detail the overlapping damages in construction images, facilitating the analysis of damages. The open-source damage classification (dacl) dataset for semantic bridge damage inspections was filtered and modified according to a predefined criterion to include textual captions, enabling fine-tuning of the BLIP and BLIP2 models. The parameters considered for captioning the images were the location and nature of the damage. The developed application reduces the time and human effort required for safety inspections of concrete structures.

Keywords: Image-captioning, Deep Learning, Structural damages

DOI: 10.18154/RWTH-CONV-254876. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction and Problem Statement

Artificial Intelligence (AI) has been adopted in the construction field for various applications, including site monitoring and performance evaluation via robotics and drones, tender evaluation, conflict resolution, sustainability assessments, waste management, safety, and hazard management [1]. Damage inspection is a key component of Structural Health Monitoring (SHM), used across various domains such as civil, mechanical, aerospace, and energy infrastructure, to evaluate structural integrity without physical intervention. It covers the detection of cracks, corrosion, wear, spalling, and other anomalies, often via non-destructive testing (NDT) methods such as ultrasonic, acoustic, magnetic flux leakage, and visual and optical techniques. Damage inspection has traditionally been human-

driven and, therefore, requires time and is prone to human error. The deployment of AI applications accelerates and improves safety inspections by automating the workflows, thereby reducing human dependency and reducing inspection errors and delays.

AI, specifically Computer Vision (CV), has been widely used to create applications for safety monitoring, progress tracking, productivity measurement, and quality control [2]. Construction damages in reinforced concrete are often complex, with different overlapping damages. This makes it challenging for CV-based models to detect and segment these damages, resulting in complex and cluttered detections or segmentations. In the case of semantic segmentations of reinforced concrete images, the model often prioritizes one overlapping damage class over the other, resulting in an unclear understanding of the structural stability from the image. This paper proposes an alternative methodology by using textual captions generated by an image captioning model to explain the context of construction damage in an image, thereby eliminating the need for human intervention and enhancing the efficiency of safety inspections.

2 State of the Art: Captioning Damages to explain semantic context

Deep learning-based computer vision models have been widely used to classify, detect, and semantically segment construction damages from images. An example of this is the dacl challenge for segmentation of semantic bridge damage [3]. The task of this challenge was to develop an image segmentation model using deep learning to segment images of damage on reinforced concrete bridges.

Textual descriptions explain the nature of overlapping damages in comprehensive captions, thereby enhancing the semantic understanding of the overlapping damages in reinforced concrete. Deep learning-based image captioning models have been widely applied in various fields, including medical image analysis, semantic tagging, image retrieval, and content creation. In the construction domain, an existing study investigates whether deep learning techniques can generate accurate, descriptive captions for construction site images to support scene understanding and documentation [4]. The paper titled “Manifesting construction activity scenes via image captioning” [5] introduces a novel, automated framework that generates descriptive captions of real-world construction activity scenes using image captioning techniques. The applications of Large Language Models (LLMs) in the construction domain are limited [6]. There is no existing research on the implementation of image captioning for construction damage images to explain the interrelationships between these damages.

In summary, existing AI applications in the construction domain for safety inspections primarily focus on Computer Vision to infer information from images. In contrast, the proposed approach, which uses a VLM fine-tuned on a captioning dataset, can potentially better define the context of complex and overlapping damages than semantic segmentation.

3 Methodology

Deep learning-based image captioning models aim to generate descriptive text for images automatically. These models traditionally used a Convolutional Neural Network (CNN) to extract visual features from an image, followed by a Recurrent Neural Network (RNN) to generate a sequence of words that form a caption. Recent models utilize transformer architecture, which employs a self-attention mechanism

for both the encoder and decoder, focusing on relevant image regions during the captioning process. This approach has enabled significant progress in developing more accurate and contextually relevant descriptions. Bidirectional Encoder Representations from Transformers (BERT) [7], Vision transformer (ViT) [8], Bootstrapping Language-Image Pretraining (BLIP) [9], and Bootstrapping Language-Image Pretraining 2 (BLIP 2) [10] are some examples of transformer architecture-based image captioning models.

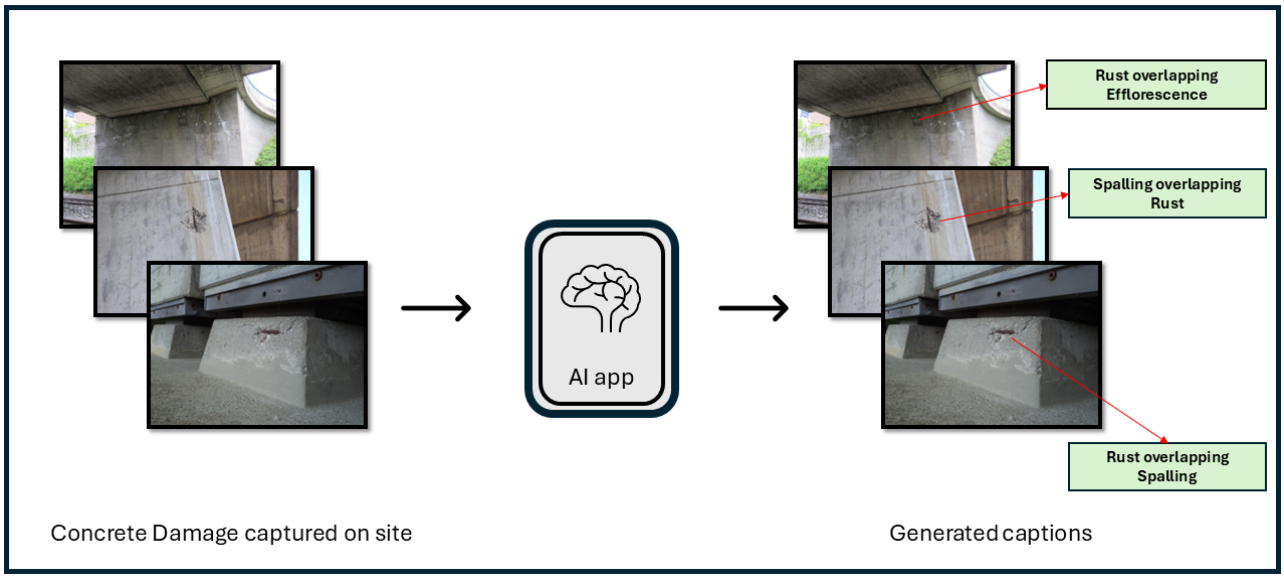


Figure 1: Proposed Methodology: The proposed approach is for a safety inspection application. The site engineer takes a picture of the reinforced concrete damage, which is then uploaded to an image captioning application. The application generates captions explaining the nature of the overlapping damages in the picture.

The BLIP architecture is particularly suitable for this project on construction damage captioning due to its ability to generate context-aware, detailed descriptions from visual input. Unlike traditional image captioning models that rely solely on paired image-text datasets, BLIP leverages both web-scale noisy data and high-quality image-text pairs through its bootstrapped training strategy, resulting in a model that generalizes well even with limited fine-tuning data. This makes it ideal for a domain like construction, where large annotated caption datasets are scarce. BLIP's architecture, as shown in Figure 2, combines a vision transformer encoder with a text decoder, enabling it to attend to fine-grained damage features, such as cracks, corrosion, or material overlaps, while generating coherent and meaningful captions. Its strong performance on multimodal tasks, combined with pre-trained checkpoints and easy adaptation, makes BLIP a robust foundation for generating captions that explain overlapping structural damages in real-world construction images.

BLIP-2 is an advanced multimodal model that improves on the original BLIP by efficiently bridging vision and language understanding. Unlike BLIP, which directly processes images and text together, BLIP-2 uses a lightweight image encoder combined with a frozen large language model (LLM), connected via a learnable query network. This design significantly reduces computational costs while maintaining strong performance on image captioning, visual question answering, and cross-modal

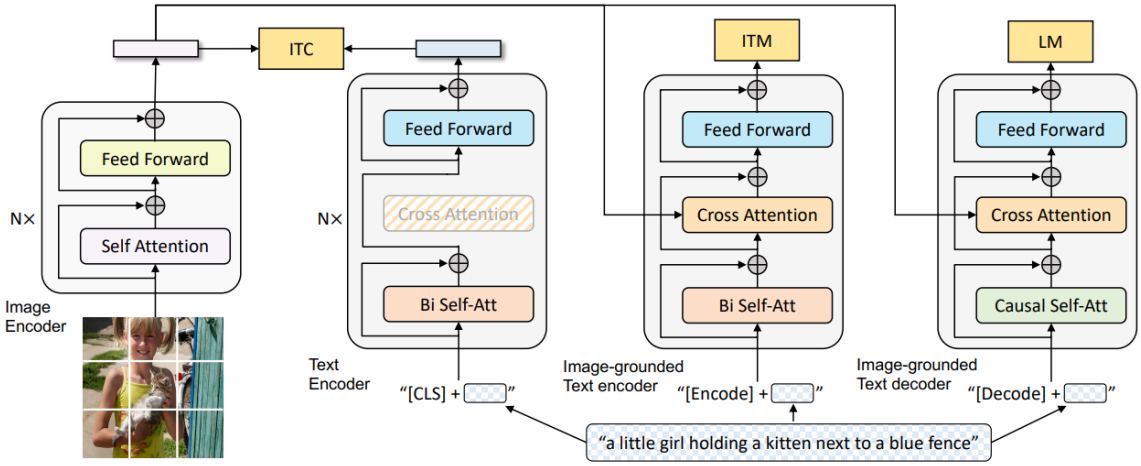


Figure 2: BLIP Architecture [9]

reasoning. Its advantages over BLIP include faster training, lower memory requirements, and the ability to leverage powerful pre-trained LLMs for richer, more coherent text generation from images.

Evaluation metrics for image captioning models assess how the generated captions align with human-written references. The fine-tuned model is evaluated using various metrics like BLEU, METEOR, and ROUGE.

BLEU (Bilingual Evaluation Understudy)
BLEU is a metric to evaluate the quality of machine-generated translations or captions against reference translations or captions. It measures the overlap between the generated and reference captions based on n-grams (continuous sequence of n words) present in both. BLEU score ranges from 0 to 1; a score of 1 means a perfect match, while a score of 0 means no overlap between the generated and reference captions. BLEU's primary focus on exact word matches often results in a lack of correlation with human judgment, especially in cases where synonyms or paraphrasing occurred [11].

METEOR (Metric for Evaluation of Translation with Explicit ORdering)
This metric considers exact, stem, synonym, and paraphrase matches, balancing precision and recall, and often better correlates with human judgment than BLEU. The METEOR score ranges from 0 to 1, with a higher score indicating better caption quality [12].

ROUGE (Recall-Oriented Understudy for Gisting Evaluation)
ROUGE focuses on recall-based n-gram and sequence matching, a technique commonly used in text summarization and applicable to captioning. ROUGE ranges from 0 to 1, with higher scores indicating greater similarity between the generated caption and the reference [13].

All of these metrics offer complementary perspectives, and combining them provides a more comprehensive picture of the captioning performance of the fine-tuned model.

4 Implementation

The project was developed in two phases. In the first phase, the dacl dataset was filtered and captioned using predefined criteria. In the second phase, BLIP and BLIP2 models have been finetuned on the overlapping damages-captions dataset, and then evaluated and deployed in a web app.

Of the 19 total available classes, eight damage classes (crack, wetspot, efflorescence, rust, rock pocket, spalling, cavity, and exposed rebar) were selected based on their impact on the structural stability of reinforced concrete bridges. The captions have been automatically generated by a custom filtering and captioning script for all training images using the polygonal annotations from the dacl dataset. The criterion of the script for filtering is to filter and retain images from the original dataset that contain the selected eight damage classes and include at least two damage classes that overlap with each other. The resulting 1,500 images were used to fine-tune the image captioning model. Additionally, 200 images were used for evaluating the fine-tuned model. The images were captioned with parameters indicating the damage classes present in the image, their locations, and whether any overlapping damage was also present. To determine the position of the damage in the image, the image is divided into nine equal sections, as shown in Figures 3 and 4. Each caption is structured with the image name, image path, and the caption itself as shown in Algorithm 1. The training caption specifies the presence of each damage instance overlapping another damage instance, and the captions of all the images are stored in a nested JSON file.

top-left	top-center	top-right
mid-left	mid-center	mid-right
bottom-left	bottom-center	bottom-right

Figure 3: Image is divided into nine sections, which are used in the captioning schema

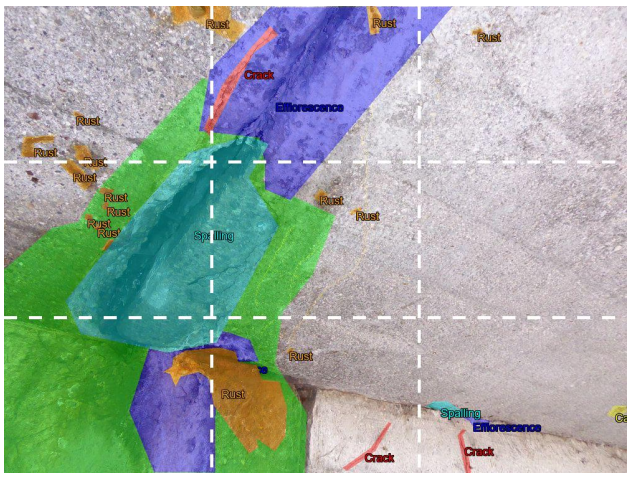


Figure 4: Example training image showcasing the various overlapping damages

Algorithm 1: Example Caption

```
{
  "image_id": "dacl10k_v2_train_0179",
  "image_path": "output/train_overlays/dacl10k_v2_train_0179.jpg",
  "caption": "wetspot overlapping efflorescence damage at mid-left; wetspot
      overlapping spalling damage at mid-left; wetspot overlapping
      efflorescence damage at mid-left; wetspot overlapping crack damage at mid
      -left; wetspot overlapping rust damage at mid-left; wetspot overlapping
```

```
        rust damage at mid-left; wetspot overlapping rust damage at mid-left;
        wetspot overlapping rust damage at mid-left; wetspot overlapping rust
        damage at mid-left; wetspot overlapping rust damage at mid-left; wetspot
        overlapping efflorescence damage at mid-left; efflorescence overlapping
        spalling damage at top-center; efflorescence overlapping crack damage at
        top-center; efflorescence overlapping rust damage at top-center; spalling
         overlapping efflorescence damage at mid-left; spalling overlapping rust
        damage at mid-left; spalling overlapping efflorescence damage at mid-left
        ; efflorescence overlapping rust damage at bottom-left."
}
```

The BLIP model was fine-tuned on the fine-tuning dataset using a batch size of 2 and a learning rate of 5e-5, while the BLIP2 model was trained with the same batch size and learning rate of 10e-5. The developed model was then evaluated and deployed in a web-based application using Gradio, which is an open-source Python library for quickly creating an interactive frontend.

5 Results

Both the BLIP and BLIP2 models fine-tune quickly on the relatively small dataset. Training and validation losses decrease exponentially and then stabilize after a few epochs. This indicates a good fit of the model to the training data. It is also observed that when fine-tuned for longer epochs (>20), the model begins to overfit to the dataset, as observed by the slightly increasing validation loss after epoch 30.

Table 1: Image Captioning Model Evaluation Metrics indicate better scores for BLIP2 model as compared to the BLIP model.

Sr. No.	Model	BLEU	METEOR	ROUGE
1	BLIP	0.19	0.17	0.21
2	BLIP2	0.23	0.45	0.27

The fine-tuned models are evaluated on the validation dataset of 200 images. The generated captions are stored in a nested JSON file, and the average BLEU, METEOR, and ROUGE scores are computed for the models. The scores are tabulated below in Table 1. The results indicate that the BLIP2 model performs better than the BLIP model. It is observed that the captions generated by BLIP2 are in the correct schema and coherent as compared to the BLIP captions. Therefore, the BLIP2 model has learned to recognize damage in concrete images from captions and can describe them on new, unseen images to some extent. The BLIP2 model sometimes generates inaccurate or repetitive captions. Hence, there is still some room for improvement in the caption generation capability of the model, which is discussed in section 7. The BLIP2 model is then deployed in the web application using Gradio as seen in Figure 5.

6 Conclusion

This paper proposed the development of a novel image captioning model for explaining overlapping concrete damages. A captioning dataset was constructed according to predefined criteria from the

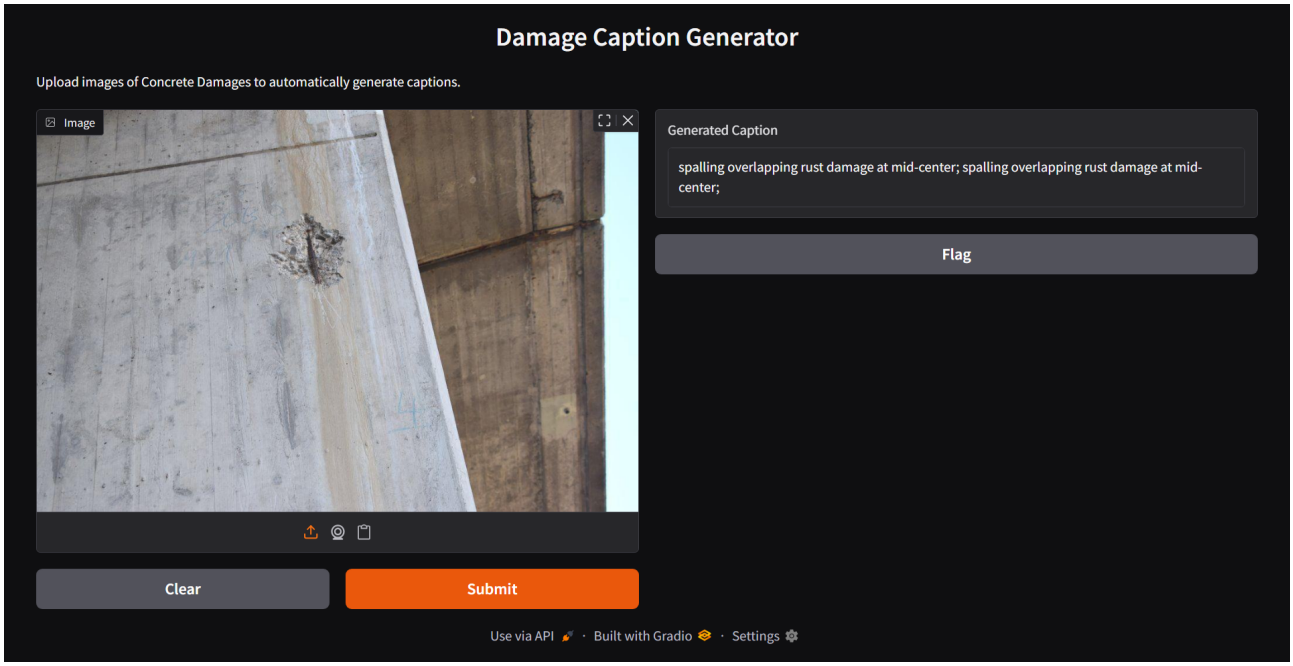


Figure 5: Damage Captioning Generator: Upload an image with concrete damage to the app, and it generates the corresponding caption as illustrated.

dacl dataset, and a structured captioning schema was developed to annotate the data systematically. This dataset was utilized to fine-tune BLIP and BLIP2 models, which were then evaluated using various metrics, yielding promising initial results. The better-performing BLIP2 model was then deployed in a local web application to generate captions explaining the damage in the images. Currently, the model performs satisfactorily, although it lacks the precision to be used in a commercial application. With further fine-tuning on a better captioned dataset, the model can give improved results.

This application significantly reduces the time and effort required for real-world safety inspections. It automates the inspection workflows and minimizes the errors in inspection checks.

7 Future Work

The model's performance can be further improved by utilizing a more effective captioning schema and a more advanced model, such as GPT2 (Generative Pre-Trained Transformers), or by employing a CNN-RNN architecture-based model. The captioning schema can be extended to include additional parameters, such as a description of the severity of the damage. In this way, the captioning dataset for the BLIP model can be further refined to produce better captions. Therefore, there is still substantial scope for improving the captioning schema and developing optimized fine-tuning strategies for the model to enhance the accuracy of the generated captions.

In this project, the model has been deployed in a web-based application to automatically and efficiently describe the context between overlapping damages in reinforced concrete structures, thereby reducing the time and effort required for the task. The generated captions can be further stored in a database and linked to structural elements in a digital twin or IFC model of the concrete structure. This application

could additionally be used for automatic image captioning to annotate pictures of reinforced concrete structures, creating a new caption dataset.

Data availability statement

The data supporting this study's findings are available at https://gitlab.lrz.de/000000003B9CC712/fbi-2025-image-captioning-concrete-damages, ensuring that they are Findable, Accessible, Interoperable, and Reusable (FAIR).

Acknowledgments

The authors gratefully acknowledge the support of the Nemetschek Innovation Foundation as part of the AInspect2Twin project, enabling our research and its practical application.

References

[1] S. O. Abioye et al., "Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges", *Journal of Building Engineering*, vol. 44, p. 103 299, 2021. DOI: https://doi.org/10.1016/j.jobe.2021.103299 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352710221011578

[2] S. Paneru and I. Jeelani, "Computer vision applications in construction: Current state, opportunities & challenges", *Automation in Construction*, vol. 132, p. 103 940, 2021. DOI: https://doi.org/10.1016/j.autcon.2021.103940 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580521003915

[3] J. Flotzinger, P. J. Rösch, and T. Braml, *Dacl10k: Benchmark for semantic bridge damage segmentation*, 2023. arXiv: 2309.00460 [cs.CV]. [Online]. Available: https://arxiv.org/abs/2309.00460

[4] B. Xiao, Y. Wang, and S.-C. Kang, "Deep learning image captioning in construction management: A feasibility study", *Journal of Construction Engineering and Management*, vol. 148, no. 7, p. 04 022 049, 2022.

[5] H. Liu, G. Wang, T. Huang, P. He, M. Skitmore, and X. Luo, "Manifesting construction activity scenes via image captioning", *Automation in Construction*, vol. 119, p. 103 334, 2020. DOI: https://doi.org/10.1016/j.autcon.2020.103334 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580520309146

[6] A. Saka et al., "Gpt models in construction industry: Opportunities, limitations, and a use case validation", *Developments in the Built Environment*, vol. 17, p. 100 300, 2024. DOI: https://doi.org/10.1016/j.dibe.2023.100300 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666165923001825

[7] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, *Bert: Pre-training of deep bidirectional transformers for language understanding*, 2019. arXiv: 1810.04805 [cs.CL]. [Online]. Available: https://arxiv.org/abs/1810.04805

[8] A. Dosovitskiy et al., *An image is worth 16x16 words: Transformers for image recognition at scale*, 2021. arXiv: 2010.11929 [cs.CV]. [Online]. Available: https://arxiv.org/abs/2010.11929

[9] J. Li, D. Li, C. Xiong, and S. Hoi, *Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation*, 2022. arXiv: 2201.12086 [cs.CV]. [Online]. Available: https://arxiv.org/abs/2201.12086

[10] J. Li, D. Li, S. Savarese, and S. Hoi, *Blip-2: Bootstrapping language-image pre-training with frozen image encoders and large language models*, 2023. arXiv: 2301.12597 [cs.CV]. [Online]. Available: https://arxiv.org/abs/2301.12597

[11] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "Bleu: A method for automatic evaluation of machine translation", in *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, ACL, 2002, pp. 311–318. [Online]. Available: https://www.aclweb.org/anthology/P02-1040.pdf

[12] S. Banerjee and A. Lavie, "Meteor: An automatic metric for mt evaluation with improved correlation with human judgments", in *Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*, 2005, pp. 65–72. [Online]. Available: https://www.cs.cmu.edu/~alavie/METEOR/meteor-1.5.pdf

[13] C.-Y. Lin, "Rouge: A package for automatic evaluation of summaries", in *Text Summarization Branches Out: Proceedings of the ACL-04 Workshop*, 2004, pp. 74–81. [Online]. Available: https://aclanthology.org/W04-1013.pdf

BIKE-SCAN: A GIS Extension for Bicycle Infrastructure Safety Analysis Based on Street-Level Images

Michael Disser[1]

[1]Institute of Numerical Methods and Informatics in Civil Engineering, Technical University of Darmstadt, Darmstadt, Germany

E-mail(s): disser@iib.tu-darmstadt.de

Abstract: Bicycle infrastructure safety assessment remains a critical challenge in urban planning, with existing approaches limited by reliance on incomplete geospatial data and lack of detailed scene understanding. This paper presents BIKE-SCAN, a novel GIS-based framework that integrates multiple data sources for comprehensive bicycle safety evaluation. The system combines OpenStreetMap data, accident statistics like “Unfallatlas”, crowdsourced sensor measurements, and street-level imagery analysis to assess cycling infrastructure safety at city scale. The methodology employs computer vision techniques including trained CNNs for semantic segmentation (using Mapillary Vistas dataset) and KITTI depth estimation to extract geometric parameters from street-level images. A weighted evaluation framework incorporating nine safety parameters including accident statistics, road conditions, lane arrangements, and junction designs enables systematic safety scoring for each junction and street edge. The system is implemented as an ArcGIS extension to ensure integration with existing planning workflows. A proof of concept based on Darmstadt demonstrates reliable identification of traffic situations and infrastructure conditions, with the application designed to be generally applicable to other cities.

Keywords: Safety-Analysis, Computer-Vision, Geographic Information System, Geoinformatics

DOI: 10.18154/RWTH-CONV-254877. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

The reduction of CO_2 and particulate emissions, health benefits, and the reduction of traffic jams in city centers are some of the advantages that promote cycling [1, p. 7ff]. However, in many German cities, ensuring bicycle safety remains a significant challenge, despite ongoing political efforts to reorganize urban traffic spaces [2]. While the “German National Cycling Plan 3.0” aspires to make cycling both more appealing and safer, accident statistics indicate that bicycle-related incidents are not noticeably declining [3]. While the “STOP-principle” (substitution, technical-, organisational-, and personal safety measures) for accident prevention in cycling mentions measures such as personal protective equipment (helmets) and safety training, these measures are far less efficient than isolating

bicycle traffic through structural separation and eliminating contact points between cyclists and motor vehicles [4, p. 118]. In addition to objective safety measurements, subjective safety plays a major role in promoting cycling: While 42% of Germans consider cycling in their city too dangerous, only 17% of Dutch respondents share this concern [5]. This difference can be attributed to the Netherlands' 28% share of cycle paths (compared to 12% in Germany) [1, p. 11] and structural design rules that physically separate cycle paths on both roads and junctions ("Dutch system") [4, p. 54ff].

To make cities more bicycle-friendly, problematic areas in the transport network must be identified and improvement suggestions incorporated. Existing research on safety-critical location analysis (see Section 2) often uses OpenStreetMap data to identify danger spots. However, this data lacks information about detailed traffic conditions, such as parked car positioning, and therefore does not include dooring risks or insufficient overtaking distances. This work addresses these limitations by developing BIKE-SCAN, a geographic information system (GIS)-based framework that combines street-level imagery evaluation with computer vision methods to analyze detailed traffic situations. Through this analysis, combined with existing geo-referenced safety statistics, safety-critical locations in the transport network can be identified. Risk factors are derived separately for road edges and junctions from specialist literature and regulations, with each receiving a bicycle safety score. For example, structurally separated cycle paths receive better safety scores than mixed traffic, and the "Dutch system" at junctions rates better than unprotected cycle lanes. To systematically address these challenges, this research investigates three key questions: **RQ1**: Which geometric and semantic factors in road infrastructure influence bicycle safety? **RQ2**: Which (open) data sources enable robust safety analysis? How can street-level imagery improve this analysis? **RQ3**: How can unsafe areas be reliably identified on a city-wide level?

The remainder of this paper is organized as follows: In Section 2, related work in cycling safety assessment is reviewed; Section 3 defines the safety parameters and evaluation framework; Section 4 presents the BIKE-SCAN methodology and implementation; Section 5 discusses results from the Darmstadt proof of concept; and Section 6 concludes with future research directions.

2 Background and Related Work

Cycling infrastructure safety assessment has shifted from traditional engineering guidelines to data-driven methods over recent decades. This section reviews the current state of research in cycling safety assessment, focusing on the integration of geospatial data, computer vision, and safety analysis.

The foundation of cycling safety assessment lies in established engineering guidelines and design principles. German construction standards, including the "Richtlinien für die Anlage von Stadtstraßen" (RAL) and "Empfehlungen für Radverkehrsanlagen" (ERA), provide comprehensive frameworks for safe cycling infrastructure design. Guidelines like [1], [6], and [4] emphasize bike-friendly urban planning and highlight knowledge such as the importance of physical separation between cyclists and motorized traffic, with the "Dutch system" representing the gold standard for intersection design. The integration of GIS with cycling safety analysis has enabled large-scale, systematic assessment of urban cycling networks. The potential of OpenStreetMap (OSM) data for safety assessment has been demonstrated in [7], [8]. While [7] generated safety scores for cycling routes in Bonn using

multivariate linear regression, [8] developed cross-city models predicting accident severity across London, Boston, and Pittsburgh. Both studies confirmed that protected cycle lanes significantly improve cyclist safety and validated the potential for data-driven safety assessment. [9] highlighted significant gaps in network-level bicycle data assessment, noting that OSM data often lacks detailed infrastructure information and suffers from topological inconsistencies, restricting the reliability of purely OSM-based safety assessments.

The advent of street-level imagery and computer vision techniques has opened new possibilities for detailed infrastructure assessment. [10] developed a bikeability index using Google Street View imagery and computer vision, creating 34 indicators to assess cycling conditions in Singapore and Tokyo. Their results demonstrated that street-level imagery indicators outperformed traditional geospatial approaches in assessing urban bikeability. [11] developed methods for classifying bicycle infrastructure using street-level images captured from bicycles, providing insights into infrastructure quality from the cyclist's perspective. The integration of sensor data and crowdsourced information has enhanced the granularity of cycling safety assessment. OpenBikeSensor projects collect overtaking distance measurements from cyclists, providing real-world data on vehicle-cyclist interactions [12]. The INFRASense project represents an approach to multi-source data integration, combining accident data, complaint systems, AI-powered damage analysis, and traffic volume measurements [13].

Most existing approaches rely heavily on single data sources, focusing primarily on either infrastructure characteristics or accident statistics. The integration of multiple data streams, including sensor data, street-level imagery, georeferenced accident statistics like "Unfallatlas", and traditional geospatial information, remains mostly unexplored. Furthermore, existing computer vision applications in cycling safety assessment have primarily focused on object detection and basic infrastructure classification. The potential for detailed geometric analysis, including measurements of lane widths, parking distances, and physical separations, has not been fully realized. Current methodologies also lack integration with practical urban planning workflows, as most research tools are designed for academic analysis rather than operational use by urban planners. This work addresses these research gaps by developing BIKE-SCAN, an integrated approach that combines GIS-based analysis with advanced computer vision techniques. The system integrates multiple data sources, including OSM data, accident statistics, crowdsourced sensor measurements, and street-level imagery, into a unified safety assessment framework, creating scene understanding capabilities for sections of the city. Furthermore, BIKE-SCAN is designed as an ArcGIS extension, ensuring integration with existing urban planning workflows.

3 Safety Parameters for Analyzing Bike-Friendly Infrastructure

Based on literature research of construction guidelines ("Richtlinien für die Anlage von Stadtstraßen", "Empfehlungen für Radverkehrsanlagen") and literature on bike-friendly construction [1], [4], [6], the most relevant parameters for objective and subjective cyclists' safety have been developed and summarized in Table 1. The parameters are based on data sources that will be utilized in the BIKE-SCAN workflow. They contain criteria for calculating a specific safety score in the risk assessment. To divide the road network into assessable areas, this analysis is carried out for junctions and the road edges between the junctions, which may contain one or more lanes in each direction.

Table 1: Safety Parameters and Evaluation Framework for Bicycle Infrastructure Analysis.

Parameter (Weight)	**Data Source (Evaluation Criteria)**	**Safety Score (Low Scores Indicate Safer Conditions)**
Common Parameters for Junctions and Edges		
Accident Statistics (0.20)	Geodata: Unfallatlas (bicycle-involved, last 5 years)	1: none, 2: minor injury only, 3: one serious injury or multiple minor, 4: deadly accident or multiple injuries, 5: multiple accidents including deadly
Overtaking Distance (0.15)	Geodata: OpenBikeSensor (average measurement)	1: >1.5m, 2: 1.0-1.5m, 3: 0.8-1.0m, 4: 0.5-0.8m, 5: <0.5m
Road Condition (0.05)	Geodata: OSM (attributes: surface, smoothness, condition)	1: asphalt/concrete, excellent/good smoothness, 2: paved, intermediate smoothness, 3: gravel/cobblestone, bad smoothness, 4: dirt/sand, very bad smoothness, 5: mud, horrible smoothness, damaged condition
Bike Lane Visibility (0.10)	Image Analysis (visibility, markings)	1: visible (bright green/red), 2: visible markings, 3: no coloring, thick line, 4: no coloring, sign only, 5: none, mixed traffic
Junction-Specific Parameters		
Junction Design (0.30)	Image Analysis (junction type classification)	1: protected, separated lanes (“Dutch system”), 2: unprotected lane in middle, right turning lane to the right, 3: unprotected lanes, 4: mixed traffic, 5: mixed traffic, cycle path ended before junction
Crossing Safety (0.20)	Image Analysis (visibility, right-turn conflicts, protection)	1: fully protected, 2: partially protected, 3: marked crossing, 4: unmarked, 5: conflicting markings
Edge-Specific Parameters		
Lane Arrangement (0.25)	Image Analysis (relation of lanes, cycle paths, curbs, grass strips, barriers)	1: complete separation with curb or grass, 2: partial separation, in road, 3: marked lane, 4: separation with single white line, 5: no separation
Overtaking Feasibility (0.20)	Image Analysis (parked cars, lane width, arrangement)	1: safe, 2: possible, 3: difficult, 4: impossible, 5: dangerous
Lane Width Compliance (0.05)	Image Analysis (measured vs. required lane width)	1: exceeds standard, 2: meets standard, 3: slightly below, 4: significantly below, 5: dangerously narrow

4 BIKE-SCAN: A GIS-Analysis Tool for Bicycle Infrastructure

To combine a GIS-based approach using street-level imagery analysis and georeferenced safety-related data, the application BIKE-SCAN (Bicycle Infrastructure Knowledge-based Evaluation using Safety-checking and Crowd-sourced knowledge for an Analysis Network) was developed. The application combines OSM analysis with additional georeferenced safety data such as accident statistics and enhanced scene understanding based on street-level imagery. The tool is designed for easy use within existing GIS planning environments and can be expanded by adding further georeferenced information or creating new safety analysis rules based on image interpretation. The application combines GIS techniques for handling georeferenced data with a Python-based extension for image evaluation. Figure 1 shows an overview of the BIKE-SCAN workflow components.

4.1 GIS-Extension

The process begins by importing georeferenced safety data (*OpenBikeSensor* [12] and relevant *Unfallatlas* [14] data into ArcGIS as shapefiles). OSM road data is then imported and filtered for relevant objects including highways, junctions, cycle paths, and street signs. Due to inconsistent

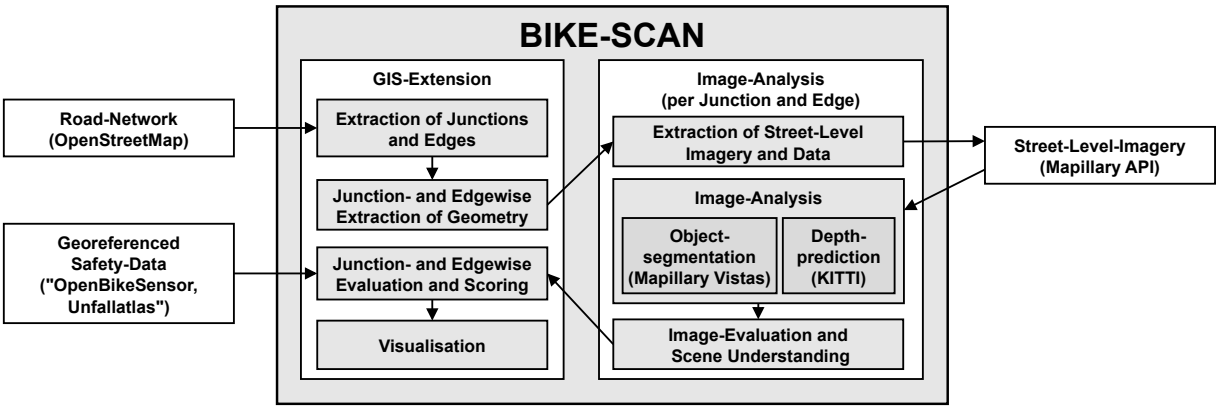


Figure 1: Overview of the Workflow of the BIKE-SCAN Application.

junction marking in OSM, junctions are identified using “Spatial Join” operations to locate where three or more highway instances meet. A junction object with a set radius (user-configurable, e.g., 30 meters) is placed to cover all relevant information. Since OSM road edges contain only center polylines and lane numbers (often missing width information), edges receive an extended feature class. This follows the road polylines with a width encompassing all edges, including cycle paths and sidewalks beside the road edge. The width is extended by several meters to include images and georeferenced safety data that may not be precisely located. After creating junctions and edges, each instance is analyzed for geometry and information such as lane numbers and outlines. This information triggers the *Image-Analysis* process.

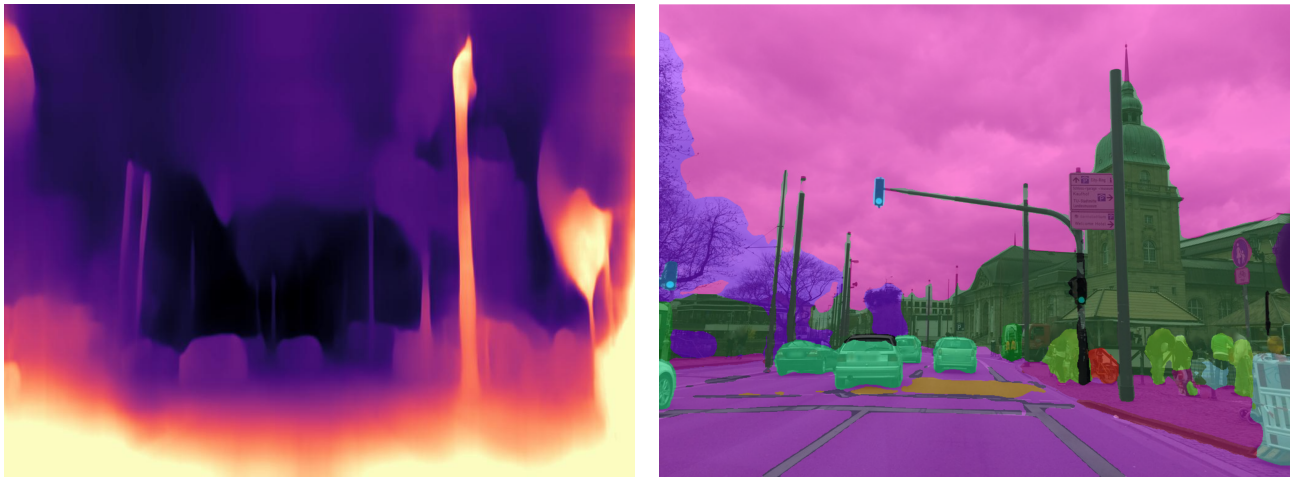


Figure 2: Image Analysis: Depth Prediction with KITTI (Left) and Object Segmentation with Mapillary Vistas (Right). Original Image by Mapillary-User with ID 100474358864037.

4.2 Image Analysis

Street-level imagery of the desired area is requested through structured API queries of the Mapillary API. For edges, each edge is divided along its polyline into segments of specified length (e.g., 20 meters). A defined number of images (e.g., 5) are requested for each segment using a filter set. The filter queries images that align with the polyline direction (looking in the travel direction), excludes

pictures older than a specified time period and panorama images, removes images without intrinsic camera parameters, and filters out images angled too far toward the sky or ground. This selection process obtains suitable images without excessive requests to the Mapillary API. Images for junctions are selected similarly, with a specified number of images for each edge connected to the junction, directed toward the junction center.

The saved images are evaluated individually in the image analysis process. Pixel-by-pixel object segmentation is performed using Mask2Former convolutional neural network (CNN) with Swin Transformer backbone trained on the Mapillary Vistas dataset, providing 65+ street-level object classes including lane markings, curbs, vehicles, and traffic infrastructure. Distance determination is carried out using KITTI Depth Estimation based on ResNet-18 encoder with custom depth decoder (Monodepth2) to obtain pixel-by-pixel distance estimates, as shown in Figure 2. Results including class types of recognized instances, their pixel-by-pixel image locations, and depth information are saved in JSON format and passed to the next step.

4.3 Image Evaluation, Scene Understanding and Scoring

Based on the *Image-Analysis* results and GIS-Extension input, edge sections and junction directions are examined. Using object segmentation results, lanes separated by curbs, lane markings, or parking cars are identified from the right side to the road center. Lane geometry is combined with depth estimation to measure lane widths. This involves selecting pixels at the left and right lane boundaries at the same depth (e.g., 15 meters) and creating a horizontal line. Using real-world depth values, camera intrinsic parameters, and pinhole camera assumptions with flat surface estimation, lane width is calculated through inverse projection. This enables street-level measurements and scene understanding based on object segmentation. A rule-based script analyzes detected object geometric relationships to achieve scene understanding, including checking for cycle path presence and physical separation by curbs, grass, or other elements. Based on this analysis, image-based parameters from Table 1 are determined, including crossing type identification, verification of legally required overtaking distances, and identification of situations causing inadequate overtaking clearance or dooring risks based on the identification of objects, the lane structure, and the comparison with the guidelines. Another example is the parameter “Bike Lane Visibility”, determined from the RGB-value of identified cycle paths, distinguishing between highly visible red/green markings and less visible gray surfaces.

Based on *Image-Evaluation and Scene Understanding* results and georeferenced safety data values, the safety evaluation framework from Table 1 is applied. Each parameter receives a specific score (low scores indicate safer conditions), and a weighted average score is calculated for each junction and edge. Results are visualized on the GIS map with relevant information explaining the scoring rationale. The identified junction and edge arrangements are also displayed for easy overview.

5 Results and Discussion

A case study for BIKE-SCAN was made using the city of Darmstadt. For the georeferenced safety data, the Darmstadt OpenBikeSensor dataset and local accident data from the Unfallatlas were imported. Initially, larger road junctions (beyond residential type) were prioritized, as these locations often had incomplete Mapillary image coverage and, according to accident statistics analysis, the majority of

bicycle accidents occurred on multi-lane roads and larger, more complex junctions. To validate the accuracy of lane width measurements, random samples of image evaluations were compared with on-site measurements. While some images enabled reliable measurements, meaningful data could not be generated from certain images due to poor image conditions and inaccurate identification of lane separations. More homogeneous data through controlled image capture (consistent camera specifications and angles) would significantly improve measurement precision. For the identification of *Junction Design*, *Lane Condition*, and *Lane Arrangement* parameters (see Table 1), traffic situations could be determined with reasonable reliability when compared to satellite imagery validation. However, separated cycle lanes were occasionally missed when obscured by parked vehicles. Cycle paths separated only by white lane markings were partially recognized, as pictograms or signs indicating cycle path designation were not visible in the available street-level images. Analysis of the results revealed that *Road Condition* parameters in OSM were set only sporadically, indicating potential for improvement through surface condition analysis approaches such as those proposed by [11].

6 Conclusion and Outlook

This work introduced BIKE-SCAN, a comprehensive GIS-based approach for systematic bicycle infrastructure safety assessment that integrates multiple data sources including OpenStreetMap data, accident statistics, crowdsourced sensor measurements, and street-level imagery analysis. Built as an ArcGIS extension (for Mapillary-Addin-Integration, otherwise using open source libraries), it integrates seamlessly into existing urban planning workflows, avoiding the creation of isolated solutions and allowing scalability and transferability to other cities and data sources. Regarding **RQ1**, the analysis identified nine critical parameters spanning accident statistics, overtaking distances, road conditions, lane arrangements, and junction designs. The weighted evaluation framework demonstrates that physical separation and protected junction designs significantly impact cyclist safety, confirming findings from cycling infrastructure research. For **RQ2**, BIKE-SCAN successfully combines publicly available, easily accessible data sources. The integration of OpenBikeSensor measurements, Unfallatlas accident data, and Mapillary imagery with computer vision analysis (object segmentation and depth/lane width estimation) enables detailed geometric measurements and scene understanding previously impossible with traditional GIS approaches alone. Addressing **RQ3**, the Darmstadt case study demonstrates that BIKE-SCAN can reliably identify traffic situations and safety conditions. Junction design, lane conditions, and lane arrangements were determined with high accuracy when compared to satellite data, though limitations exist with occluded infrastructure and sparse image coverage in residential areas.

Integration with existing lane models and municipal traffic data could provide more accurate baseline data while eliminating image query limitations and ensuring more homogeneous imagery. The incorporation of additional safety factors extractable from images presents significant potential, including subjective safety assessments, traffic guidance clarity, and perceptual load reduction. Future development should explore automated improvement recommendation systems, such as CNN training with accident data for parameter weighting optimization. Integration of additional crowdsourced data sources such as municipal complaint systems as well as including surface condition analysis, following

the INFRASense approach [13] would also improve the system. These enhancements would establish BIKE-SCAN as a foundation for evidence-based cycling infrastructure planning and improvement.

References

[1] M. Meschik, *Planungshandbuch Radverkehr*. Springer Vienna, 2008.

[2] B. für Digitales und Verkehr (BMDV). “BMDV - nationaler radverkehrsplan 3.0”. [Online]. Available: https://bmdv.bund.de/SharedDocs/DE/Artikel/StV/Radverkehr/nationaler-radverkehrsplan-3-0.html

[3] B. Impey, “Anzahl der Verunglückten bei Straßenverkehrsunfällen in Deutschland nach Art der Beteiligung von 1980 bis 2023”, Statistisches Bundesamt, Jul. 5, 2024.

[4] T. Schwendy, *Sichere Kreuzungen* (Die Blaue Reihe - Praxiswissen kompakt Band 3). Fahrradverlag, 2023.

[5] Ipsos. “Cycling across the world - global advisor survey”. [Online]. Available: https://www.ipsos.com/sites/default/files/ct/news/documents/2022-05/Ipsos-PI_Cycling_2022-05-24.pdf

[6] B. Schreck, “Cycling and designing for cyclists in germany: Road safety, guidelines and research”, Oct. 22, 2024. DOI: 10.5507/tots.2017.007

[7] W. Li et al., “Safety assessment of cycling routes in urban environments”, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLVIII-4/W10-2024, pp. 125–130, May 31, 2024. DOI: 10.5194/isprs-archives-XLVIII-4-W10-2024-125-2024

[8] S. Daraei, K. Pelechrinis, and D. Quercia, “A data-driven approach for assessing biking safety in cities”, *EPJ Data Science*, vol. 10, no. 1, p. 11, Dec. 2021. DOI: 10.1140/epjds/s13688-021-00265-y

[9] A. R. Vierø, A. Vybornova, and M. Szell, “BikeDNA: A tool for bicycle infrastructure data & network assessment”, *Environment and Planning B: Urban Analytics and City Science*, vol. 51, no. 2, pp. 512–528, Feb. 2024. DOI: 10.1177/23998083231184471

[10] K. Ito and F. Biljecki, “Assessing bikeability with street view imagery and computer vision”, *Transportation Research Part C: Emerging Technologies*, vol. 132, p. 103 371, Nov. 2021. DOI: 10.1016/j.trc.2021.103371

[11] K. Backman, B. Beck, and D. Kulić, “Classifying bicycle infrastructure using on-bike street-level images”, in *2024 IEEE ITSC*, pp. 3166–3173. DOI: 10.1109/ITSC58415.2024.10920149

[12] A. D.-D. e.V. “Map - OpenBikeSensor portal”. [Online]. Available: https://obs.adfc-darmstadt.de/map

[13] “INFRASense: Auf Fahrradstrecken aggregierte Kennzahlen zur Verkehrssicherheit”. [Online]. Available: https://www.govdata.de/suche/daten/infrasense-auf-fahrradstrecken-aggregierte-kennzahlen-zur-verkehrssicherheit-oberflachenbeschaf

[14] S. Ä. des Bundes und der Länder. “Unfallatlas | kartenanwendung”. [Online]. Available: https://unfallatlas.statistikportal.de/

End-to-End Parametric Workflow for Bridge Design with Automated Cost Analysis and Generative Modeling

Jonas Wiederer[1], Simon Hoeng[1], Maik Görgner[1,2] and Anna-Maria Mehringer[1,2]
[1]Ostbayerische Technische Hochschule Regensburg, Seybothstraße 2, 93053 Regensburg, Germany,
[2]Firmengruppe Max Bögl, Max-Bögl-Straße 1, 92369 Sengenthal, Germany
E-mail(s): jonas.wiederer@oth-regensburg.de, simon.hoeng@oth-regensburg.de,
maik.goergner@oth-regensburg.de, anna-maria.mehringer@oth-regensburg.de

Abstract: This paper presents an end-to-end workflow for conceptual bridge design that integrates parametric modeling, automated cost analysis, and data-driven evaluation. At its core lies a top-down parametric BIM skeleton that governs geometry, earthwork volumes, and adaptive excavation pits. Through automated IFC exchange and integration with cost-analysis software via an application programming interface, component volumes are calculated and linked to a template bill of quantities via matchkeys, ensuring consistent and reproducible cost estimation. The workflow not only enables reliable analysis of bridge variants based on Key Performance Indicators (KPIs), but also produces structured synthetic datasets suitable for training machine learning models. In particular, a Conditional Variational Autoencoder is conceptually introduced to demonstrate how such data can support both forward performance prediction and inverse design conditioned on KPI targets. By improving data fidelity, the approach reduces the demand for extensive training datasets while increasing the robustness of model predictions. This research therefore lays the groundwork for bridging parametric modeling and generative Artificial Intelligence, offering a scalable pathway toward more efficient and data-driven infrastructure planning.

Keywords: Building Information Modeling (BIM), Parametric Modeling, Artificial Intelligence (AI), Conditional Variational Autoencoder (CVAE), Application Programming Interface (API)

DOI: 10.18154/RWTH-CONV-254878. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Motivation

In 2024, the construction sector in Germany generated approximately €215 billion in gross value added [1]. A substantial share of this investment was directed toward infrastructure development, with bridge construction playing a critical role. Bridges are essential components of transportation networks, facilitating mobility across natural and built barriers and ensuring the continuity of road and rail systems. Many existing structures, particularly those built in the 1960s and 1970s, no longer meet modern structural or economic standards. As a result, reconstruction has become more common than

renovation. Of the 42,300 bridges on German federal roads, nearly two-thirds have a span of less than 30 meters [2]. Despite similar design principles, each project still requires individual planning, limiting scalability and consuming resources. Although standardized bridge types exist, their potential for repetition is underutilized. Manual design efforts remain high, even when only a few parameters differ. Tools such as Building Information Modeling (BIM) and parametric modeling have improved design consistency and reusability, but they often operate in isolation, disconnected from evaluation or decision support. Artificial Intelligence (AI) offers new opportunities: By learning from previous models and planning data, it can support recurring decisions, suggest optimized configurations, and provide early, data-driven guidance. Rather than treating each bridge as a standalone task, AI-enabled systems can generalize patterns from past solutions, thereby improving planning speed and quality. Integrating such learning into digital workflows enables automation of routine infrastructure design, especially in repetitive projects. This research advances digital planning with an approach that generates and learns from design variants. By linking parametric modeling with intelligent feedback loops, the system aims to provide a scalable and efficient data-driven approach to the planning of repeatable structures.

2 State of the art

Previous work introduced a parametric bridge modeling approach with a focus on semi-automated cost analysis using matchkeys and embedded design logic within a BIM environment [3]. The model enables dynamic reconfiguration directly within the CAD environment, allowing for rapid adaptation to project-specific requirements. It incorporates cost analysis capabilities through integration with 5D BIM software. The use of matchkeys as a linkage between the 3D model and external analysis tools forms the basis for the semi-automated generation of key performance indicators (KPIs) essential for informed decision-making.

The work by Ji et al. [4] on IFC Bridge proposes a schema extension that integrates parametric geometry into the Industry Foundation Classes (IFC) standard, enabling geometric constraints and relationships essential for automated bridge design. However, IFC implementations still lack the ability to fully export parametric dependencies, which limits automation when transferring data across platforms. To address these limitations, recent research has explored software integration via Application Programming Interfaces (APIs). For example, Herlé and Blankenbach [5] present a hypermedia-driven API linking BIM, GIS, and IoT domains for federated digital twins, demonstrating improved interoperability, data flow, and automation in use cases.

More comprehensively, Pishdad and Onungwa [6] examine 5D BIM workflows, showing how the integration of BIM data with cost estimation systems enables automated and continuous cost analysis. This improves project efficiency by supporting real-time budget tracking, accurate quantity takeoffs, and streamlined payment management. Wortmann and Nannicini [7] provided an overview of black-box optimization approaches in architectural design, highlighting their suitability depending on objective function characteristics, computational effort, and optimization goals. Their work illustrates that classical optimization methods can also make valuable contributions to complex design tasks. Such methods could in principle be combined with the approaches presented here, but are not discussed further, as the focus is placed on the introduction of Conditional Variational Autoencoders (CVAEs).

Danhaive and Mueller [8] introduced a performance-conditioned generative modeling approach using CVAEs for structural design. Their method combines a sampling algorithm to generate diverse high-performing design candidates with a performance-conditioned autoencoder that maps this dataset into a low-dimensional latent space. This space is intuitive for designers to explore, enabling efficient and varied structural design options.

Balmer et al. [9] applied a CVAE framework to 18,000 synthetically generated bridge designs, sampled via Latin Hypercube Sampling and evaluated using structural simulation. Costs were approximated by multiplying material volumes with unit prices from a reference database, enabling scalable, consistent outputs for model training. The CVAE, complemented by sensitivity analysis tools, proved valuable for conceptual design, supporting both performance prediction and inverse design tasks.

Finally, Hoeng et al. [10] developed a pipeline to generate synthetic datasets from parametric BIM models, combining geometry, parameters, and structured metadata. As emphasized in the studies above, such synthetic datasets are essential for training generative models, and this work proposes a structured workflow for their automated creation.

Together, these contributions outline a growing body of research that supports automated, performance-informed infrastructure design by combining parametric modeling, open standards, API integration, synthetic data, and AI-based design space exploration.

3 Methodology

This study proposes a fully automated workflow that integrates parametric modeling, cost analysis, and structured data generation, supporting both classical KPI-based bridge design analysis and the generation of synthetic datasets that can be used for training generative design models.

A bridge model is developed in *Siemens NX* [1] using parameterized geometries and matchkeys, which allow systematic variation and semantic tagging of design elements. The creation of the parametric model in *Siemens NX* is automated through an API, enabling reproducible model generation without manual interaction. The model is connected to a 5D cost analysis environment via API, allowing automated extraction and evaluation of geometry, quantities, and cost-related data. All data transfers are managed through API endpoints to ensure reproducibility and eliminate manual processing. In addition, the integration of a bill of quantities (BoQ) within the cost analysis software enhances the mapping between geometry and unit cost structures. This setup enables the creation of a diverse dataset consisting of geometrically and economically realistic bridge designs for training CVAEs.

At this stage, the concept of the workflow is presented in its entirety, while the implementation currently covers the parametric modeling and the 5D cost analysis.

4 Concept

In conventional bridge planning, engineers manually develop and evaluate design variants based on project-specific requirements. Performing this process manually is very time-consuming, and only a limited number of variants can be explored. To accelerate and simplify this process, the following concept introduces how API interfaces can support analysis workflows and improve outcomes. This

[1] https://plm.sw.siemens.com/en-US/nx/cad-online/

approach also opens multiple additional possibilities, such as the integration of AI. Furthermore, the automated loop provides a solid foundation for generating high-quality synthetic data for training a neural network, with the possibility of performing inverse design tasks through the accessibility of KPIs. In this paper, the integration of cost and Life Cycle Assessment (LCA) is discussed, as both values rely on volume-based KPI analysis. The flow diagram in Figure 1 illustrates how the workflow operates and where an integration of AI is possible.

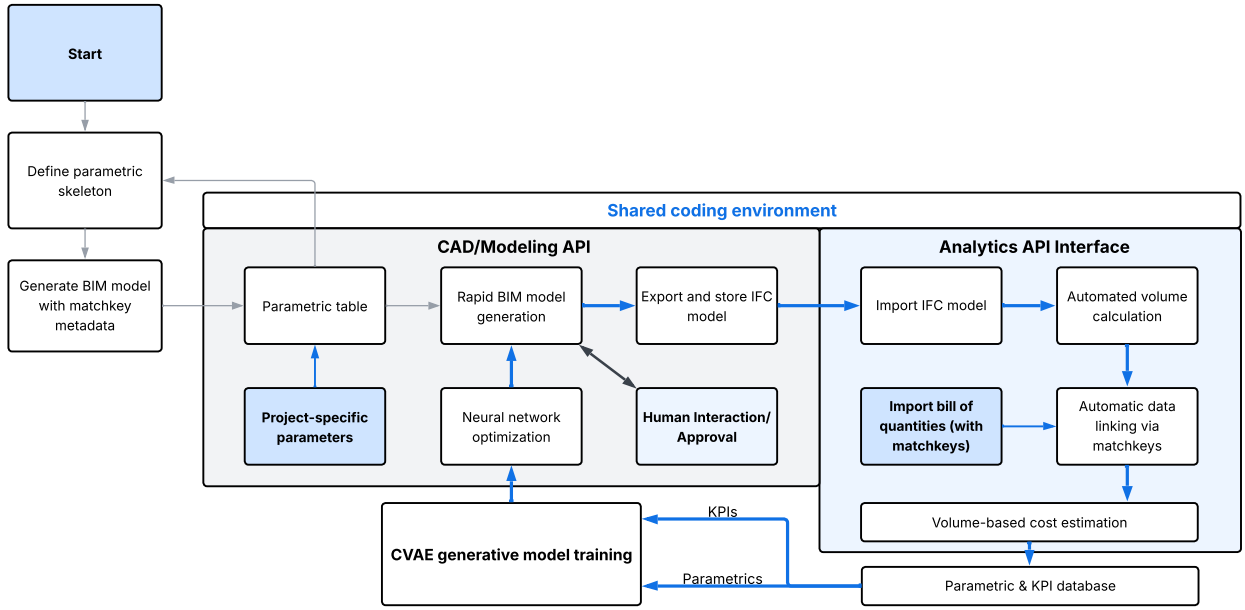


Figure 1: The workflow shows a dataflow loop built on a parametric BIM model with automated cost analysis. It demonstrates how high-quality datasets can be generated and integrated into a central database, providing the foundation for AI model training and optimization.

Starting with the parametric model, the foundation is a parametric skeleton that contains all parameters and the relationships between the individual parts. The model is built using a top-down strategy, as described in previous work, which ensures that the bridge design can be rapidly adapted. The BIM model not only includes information about the structural components themselves but also integrates metadata such as exported IFC classes and matchkey information. These matchkeys can be either uniquely generated by the designer or automatically created identifiers. By means of the parametric table, the BIM model can be fully controlled. Designers are able to integrate project-specific parameters into the CAD software and create an initial BIM model within a short time. This model can then be rapidly adapted through parametric adjustments.

Once the parametric table is defined, all subsequent steps can be centrally managed within a single coding environment. The export of the model from the CAD software and its import into the cost analysis software are executed in a fully automated manner. After importing the model into the cost analysis software, in this case *DesiteMD* [2], the volumes of the individual components are automatically calculated. In a further step, a BoQ is also imported through the interface. This document contains

[2] https://www.thinkproject.com/de/products/desite/

the matchkeys of the bridge elements, establishing the link between the positions in the BoQ and the individual components of the bridge. Although the automated import of the BoQ is not strictly necessary, it accelerates adaptation in case of changes to the document. The enriched document, which now contains cost information, can then be exported to provide the user with an initial cost estimation of the bridge.

The proposed link ensures that even small modifications in the parametric table automatically generate an updated BoQ. The workflow therefore enables highly efficient bridge analysis and opens new potential for streamlined project delivery. The generated parametric data, linked with high-quality KPI values, provides an optimal basis for producing synthetic datasets to train neural networks. In this context, a CVAE represents a particularly promising approach. It not only predicts parameter values in order to achieve specific KPI goals but also allows the definition of target KPI values from which the model can derive the corresponding parameters.

The primary goal is to develop a neural network that supports engineers in generating improved BIM models of bridges. This closes the loop, creating an efficient end-to-end workflow in which parametric modeling, cost analysis, and KPI evaluation are seamlessly integrated, thereby enabling data-driven design and optimization.

4.1 Parametric Bridge Model

In previous work [3], the structural system of the bridge was discussed in detail. Building on this, the present paper extends the scope by including the earthworks, which were previously only mentioned on a conceptual level.

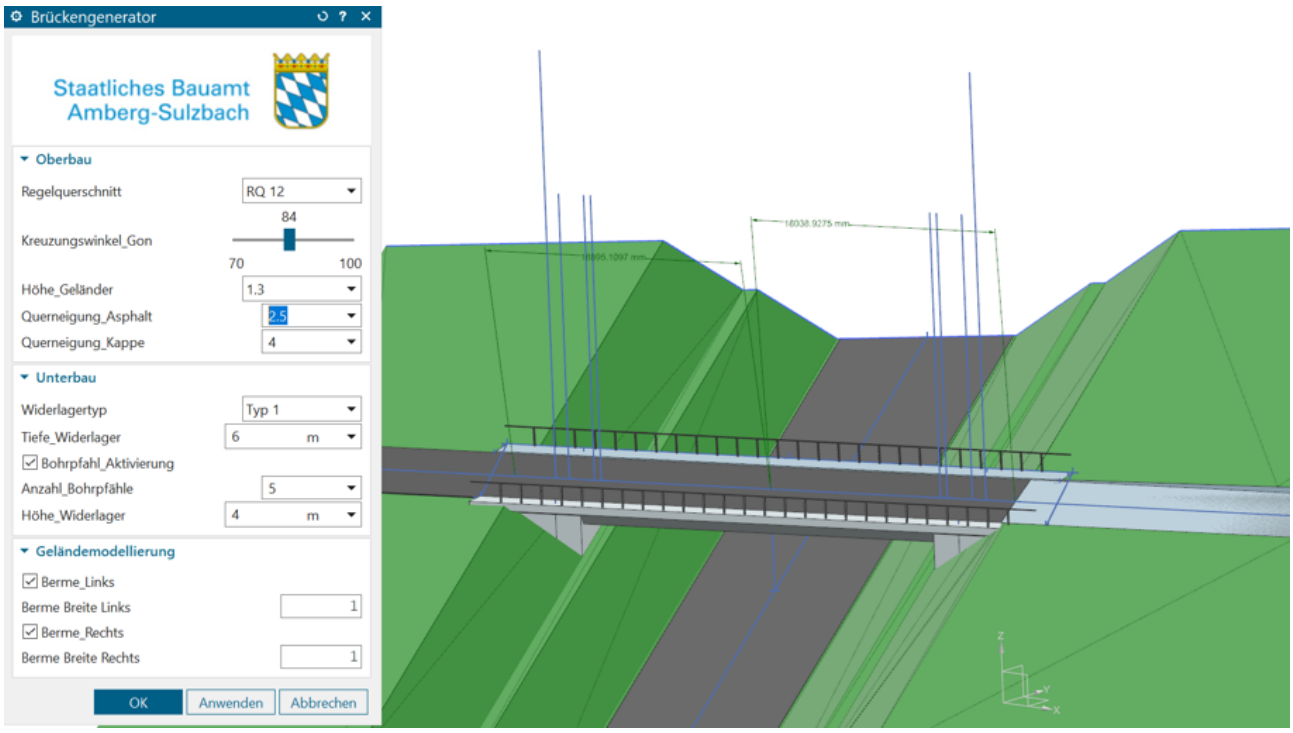


Figure 2: Bridge Generator in *Siemens NX*

As earthworks have a significant impact on the total cost, the parametric model includes an adaptive earthwork volume. The model is based on two routing lines: one for the overpassing structure and one for the underpassing structure. The overpassing line carries a terrain body, which can be modeled in a similar way to other longitudinal extrusions. The underpassing line also contains a body that cuts into the bridge structure using boolean operators. This cutout can be parametrically controlled, allowing all values and the cross-section of the earthworks to be adjusted. Furthermore, adaptive excavation pits are modeled beneath the abutments. These adapt automatically when the abutments are modified and can also affect the overall cost of the bridge. The parametric model, including the earthworks, can be fully modified via a graphical user interface named bridge generator, as illustrated in Figure 2.

4.2 Generating KPI values for a specific analysis

Through an API call, it is possible to export the model from *Siemens NX* and save the IFC file with the chosen settings, such as the IFC schema.

DesiteMD allows volume-based calculations of imported IFC models. For this purpose, a remote control interface enables sending and receiving commands from the software. The IFC bridge model is imported into *DesiteMD*, and the volumes of its components are automatically calculated. Another API call then stores these values in intermediate memory, or alternatively, the volumes can be extracted directly from the software. It is essential that, together with the volumes, the corresponding matchkey values are also requested through the call. In addition to the model import, the BoQ can also be processed. Ideally, this is based on a template BoQ that contains all possible parts, dimensions, and materials. The BoQ also includes matchkey information, enabling the linkage between the IFC model and the individual positions of the BoQ. This link is established by an internally developed script, which is triggered remotely through an API call. An into *DesiteMD* imported and automatically quantified and costed bridge model is illustrated in Figure 3.

This workflow connects the IFC model with cost calculation data and enables automated bridge cost estimation. Afterwards, the BoQ can be exported, providing external access to all results. Beyond supporting automated analysis of bridges with realistic data, the workflow can also be applied to generating synthetic datasets. By integrating an algorithm that generates random input values for the BIM model, synthetic training data can be created in the same way. This approach and its limitations will be discussed further in the discussion section.

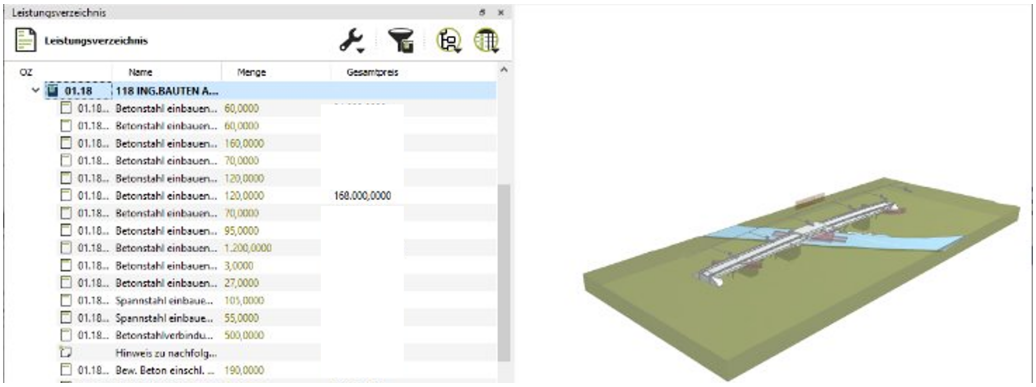


Figure 3: Calculation in *DesiteMD*

4.3 Training Data and CVAE Framework

As outlined in the workflow, the proposed methodology enables the automated evaluation of project data when sufficient real-world information is available, or alternatively, the generation of synthetic datasets. Both serve as a reliable foundation for training machine learning models. In particular, CVAEs require large amounts of training data, as emphasized by [9]. The principle is that the more abundant and higher-quality the training data, the more stable the training and the more reliable the predictions. In this study, the CVAE is introduced at a conceptual level. The training data consist of feature values, derived directly from parametric BIM model parameters and target values, which represent performance indicators such as costs and can be extended to include additional KPIs. Within the CVAE, the encoder maps the complex relationships between features and targets into a reduced latent space, where essential patterns are preserved in a compact mathematical form. The decoder then reconstructs realistic parameter–performance combinations from this latent representation.

This architecture enables forward predictions, where parameters are provided to estimate KPI outcomes, and inverse predictions, where target KPIs such as costs are specified and the model suggests suitable parameter sets. Although the detailed CVAE architecture is not elaborated in this paper, it will be the focus of future work. Here, the emphasis lies on demonstrating that the automatically generated dataset provides a robust basis for establishing such a learning framework and thereby supports the transition from traditional analysis toward data-driven generative bridge design.

5 Discussion and Limitations

The parametric bridge model forms the foundation of the entire approach. Without a consistently structured parametric model, rapid variant generation and reliable adaptations are not feasible, which also deprives the training of a neural network of a solid data basis.

We implemented end-to-end data exchange as an optimization loop. Parametrically generated variants are automatically transferred to a cost analysis application, evaluated, and the results are fed back to guide subsequent variants. This closed loop is enabled by the available interfaces of the tools involved.

A central constraint is the limited availability of reliable API interfaces across different tools. In our case, the coupling was feasible; however, for many other domain models and analyses, comparable integrations are still missing, which currently limits scalability.

From the running loop, we obtain structured, high-quality synthetic datasets. These form the basis for the conceptually employed CVAE. In the forward direction, the CVAE provides rapid KPI estimates for new parameter combinations. In the inverse direction, it generates parameter sets that meet specified target KPIs, supporting goal-directed design decisions.

Looking ahead, we will continue running the loop, curating and extending the dataset, and training the CVAE step by step. With increasing data quality, we expect reduced training effort, more realistic and robust results, and a measurable acceleration of early decision phases in bridge design. At present, throughput is primarily constrained by compute-bound runtimes on workstation hardware, particularly for variant generation, model export, and single-threaded solver runs.

Acknowledgements

This research was funded by the Bavarian Ministry of Economic Affairs, Regional Development and Energy under the Bavarian Collaborative Research Program (BayVFP) as part of the KIBIP project.

References

[1] Statistisches Bundesamt (Destatis), *Bruttoinlandsprodukt 2024: Anteile der nominalen Bruttowertschöpfung*, https://www.destatis.de, 2025.

[2] Bundesanstalt für Straßenwesen (BASt), "Brückenstatistik", Bundesanstalt für Straßenwesen, Tech. Rep., 2024. [Online]. Available: https://www.bast.de/DE/Publikationen/Statistik/Bruecken/Brueckenstatistik.pdf?__blob=publicationFile&v=2

[3] J. Wiederer, S. Höng, and F. Eder, "Konzept zur KI-gestützten parametrischen Brückenmodellierung für ressourcen- und kostenoptimierte Bauwerksentwürfe", in *35. Forum Bauinformatik*, Hamburg, Germany, 2024. DOI: 10.13154/294-10130

[4] Y. Ji, A. Borrmann, J. Beetz, and M. Obergrießer, "Exchange of Parametric Bridge Models Using a Neutral Data Format", EN, *Journal of Computing in Civil Engineering*, vol. 27, no. 6, pp. 593–606, Nov. 2013. DOI: 10.1061/(ASCE)CP.1943-5487.0000286

[5] S. Herlé and J. Blankenbach, "Hypermedia-driven RESTful API for digital twins of the built environment", en, *Automation in Construction*, vol. 165, p. 105 551, Sep. 2024. DOI: 10.1016/j.autcon.2024.105551

[6] P. Pishdad and I. O. Onungwa, "Analysis of 5D BIM for cost estimation, cost control and payments", *Journal of Information Technology in Construction (ITcon)*, vol. 29, pp. 525–548, 2024. DOI: 10.36680/j.itcon.2024.024

[7] T. Wortmann and G. Nannicini, "Introduction to Architectural Design Optimization", in *City Networks: Collaboration and Planning for Health and Sustainability*, Cham, Switzerland: Springer International Publishing, 2017, pp. 259–278. DOI: 10.1007/978-3-319-65338-9_14

[8] R. Danhaive and C. T. Mueller, "Design subspace learning: Structural design space exploration using performance-conditioned generative modeling", in *Automation in Construction*, vol. 125, 2021. DOI: 10.1016/j.autcon.2021.103588

[9] M. Balmer, P. Bucher, F. Kälin, and S. Basso, "Generative design of bridges using variational autoencoders and structural optimization", *Automation in Construction*, vol. 155, p. 105 175, 2024. DOI: 10.1016/j.autcon.2024.105411

[10] S. K. Hoeng, F. Eder, M. Schmailzl, and M. Obergrießer, "Exploring the Potential of BIM Models for Deriving Synthetic Training Data for Machine Learning Applications", in *Advances in Information Technology in Civil and Building Engineering: Proceedings of ICCCBE 2024 - Volume 2: Simulation and Automation*, 2025, pp. 54–63. DOI: 10.1007/978-3-031-87364-5_5

Automating GIS-Derived Site Plans for Digital Building-Permit

Fatma Ahmad[1] and Ralf Becker[1]
[1]Chair of Computing in Civil Engineering and Geoinformation Systems and Geodetic Institute (gia), Aachen, Germany
E-mail(s): fatma.ahmad@gia.rwth-aachen.de, ralf.becker@gia.rwth-aachen.de

Abstract: This research presents an approach to automate the generation and validation of regulatory-compliant building site plans by integrating Geospatial Information Models (GIM) and Building Information Modeling (BIM). The methodology uses Industry Foundation Classes (IFC) and CityGML to turn authoritative cadastral and zoning information into semantically rich IFC models, then exports them as XBau-compliant XML for automated permit-data exchange and rule checking. The workflow relies on precise georeferencing and boundary representation (B-Rep) modeling to define permissible building volumes constrained by the building line (Baulinie), the building boundary line (Baugrenze), and the property-line (Grundstück). The prototype employs open-source tools like ifcopenshell and demonstrates potential violations through geometric analysis. This approach streamlines building permit approval processes, reduces manual effort and errors, and advances digital transformation goals within Germany's building sector.

Keywords: BIM, Digital Building Permit, Digital Building Application, GIM, IFC

DOI: 10.18154/RWTH-CONV-254927. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

In Germany, a building permit (*Baugenehmigung*) is required before constructing, modifying, or demolishing a structure; building application (*Bauantrag*) documents—architectural drawings, structural engineering checks, energy reports and the (official) site plan (*amtlicher Lageplan*)—are still submitted as paper or PDF bundles and manually checked against zoning (*Bauplanungsrecht*) and building code (*Bauordnungsrecht*), often prompting queries and resubmissions. In 2023, 260,100 permits were approved—the lowest since 2012—with processing times of 4–8 weeks for simple cases and several months for complex ones, and a federal reform now requires all states to adopt a three-month "deemed-approved" deadline (*Genehmigungsfiktion*) by 2026. To address delays, Fauth et al. [1] call for digital transformation; since 2017, XPlanung and XBau standards have enabled structured data exchange (fully rolled out by 2022), and Mecklenburg-Western Pomerania's 2021 *digitaler Bauantrag* under the "Einer für Alle" model integrates BIM/IFC models and cadastral data for automated, paperless permitting.

Long before an architect models a building, the site plan establishes the legal and physical framework for a project by capturing parcel boundaries, existing buildings, topography, access roads, easements, utility corridors, and the proposed footprint. Survey-grade coordinate geometry ensures boundary accuracy, exposes hidden constraints, and underpins compliance checks and risk assessments. However, submitting the site plan as a flat PDF breaks the digital workflow, as automated engines require explicit geometry, topology, and semantics to verify rules for eg. “minimum clearance 3m from the property line.” By creating an open, semantically rich 3D model, this work unifies survey precision with BIM and GIS context, enabling automated compliance checking, seamless data integration, and practical application across all project stages, from site planning to city-level analysis. This paper’s novelty lies in three main contributions. The integration of BIM and GIS into a standardized 3D site-plan model, machine-readable regulatory compliance verification, and practical validation through three real-world projects.

The remainder of this paper is structured as follows. Section 2 introduces the associated project overview followed by related work and technical background. Section 5 describes the methodology of the schema transformation between CityGML and IFC on a use case. Finally, Section 6 discusses the results and Section 7 concludes and discusses the outlook direction.

2 Project Overview: *"3D-LagePlan zum Baugesuch"*

To bridge the gap between survey-grade site plans and Germany’s digital building permit workflow, the federally funded research project “3D-LagePlan zum Baugesuch” was launched within the *Zukunft Bau* programme. [2]. Over 24 months, a consortium coordinated by the *Bund der Öffentlich bestellten Vermessungsingenieure e. V. (BDVI)*—including KDS Köln, the Geodetic Institute & Chair of Computing in Civil Engineering & GIS at RWTH Aachen, the Chair of Geoinformatics at TU Munich, HHK Datentechnik, and the federal XPlanung/XBau Coordination Office will deliver Germany’s first standardized, BIM-ready 3D site-plan model for seamless integration into digital building applications.

By unifying BIM and GIM data into a single information model and automating compliance checks, the project enables planners and authorities to exchange data effortlessly and verify against XBau standards. This end-to-end digital workflow reduces paperwork, minimizes errors, accelerates approvals, and lays the foundation for smarter cities and a more responsive construction sector.

To validate and refine its prototype software in practical settings, the consortium will apply the new digital processes to three real-world projects: a school, a housing block. Feedback will be gathered to ensure the system’s robustness and versatility.

3 State of the Art

Digital permitting is advancing, but data silos, low automation, and non-machine-readable rules still hinder widespread adoption [3]. BIM–GIM integration supports zoning, yet semantic gaps and incomplete datasets persist [4], and existing rule checks often depend on manual georeferencing [5]. Recent reviews highlight the central role of georeferencing for reliable BIM–GIS integration [6], while new research stresses the need to formalize data requirements for digital permitting processes [7]. Semantic approaches for IFC–CityGML interoperability show promising results [8]. The 3-D “SitePlan ADE” demonstrates how cadastral detail can be incorporated [9]. However, a fully automated standard

workflow is still missing. We therefore introduce a lightweight tool that georeferences BIM, exchanges data via XBau, and enables surveyor-certified, rule-compliant submissions.

4 Technical Background

Geo-referencing

Geo-referencing BIM models is crucial because local coordinates in BIM tools often do not match real-world positions. IfcProjectedCRS defines the target projected coordinate system, and IfcMapConversion handles scaling, rotation, and translation from local coordinates. The transformation can be expressed as a 2D Helmert equation for horizontal coordinates:

$$\begin{bmatrix} X_{\text{UTM}} \\ Y_{\text{UTM}} \end{bmatrix} = s \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} X_{\text{local}} \\ Y_{\text{local}} \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}, \quad Z_{\text{UTM}} = Z_{\text{local}} + t_z \tag{1}$$

where s is the horizontal scale, θ the rotation angle, and t_x, t_y the translations. In equation (1), s vary based on the distance of the location from the central meridian, potentially affecting parcel boundaries or zoning compliance. In Revit, this is set via the Project Base Point and exported as IfcMapConversion and IfcProjectedCRS in IFC4+. IFC export for geo-referencing:

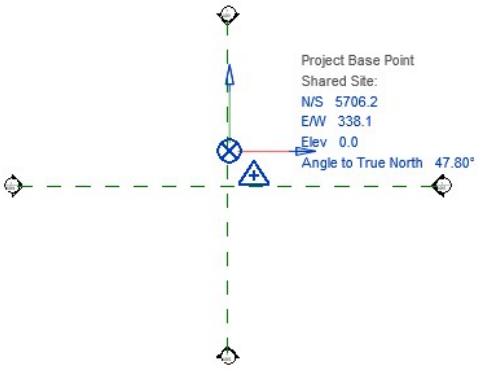


Figure 1: View aligned to Project North

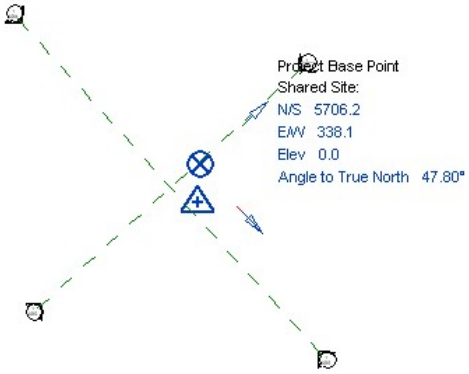


Figure 2: View aligned to True North (Revit)

```
#27 = IFCPROJECTEDCRS('EPSG:25832','ETRS89',,,,#SI_UNIT_METRE);
#28 = IFCMAPCONVERSION(#22,#27, 338.1, 5706.2, 0.0,
                       0.672708, -0.739911, 1.0);
```

Here, the rotation angle θ is encoded via its cosine and sine values: $\cos\theta = 0.672708, \sin\theta = -0.739911$. If the angle is given in degrees, it is first converted to radians before computing these values. The scale is 1.0 due to Revit export limitations.

5 Methodology

5.1 Requirement Analysis

Developing a prototype begins with a data-independent model that captures all inputs required for an official building application. Müller [10], in his master's thesis, analyzes the 3D requirements for the official site plan (*amtlicher Lageplan*) based on *BauPrüfVO NRW* §3 (24.07.2020). He categorizes object types, data sources (e.g., cadastral data (ALKIS), CityGML building models, surveyor on-site

captured data), and distinguishes between existing/planned and topographic/conceptual data—forming a solid foundation for defining 3D Lageplan inputs.

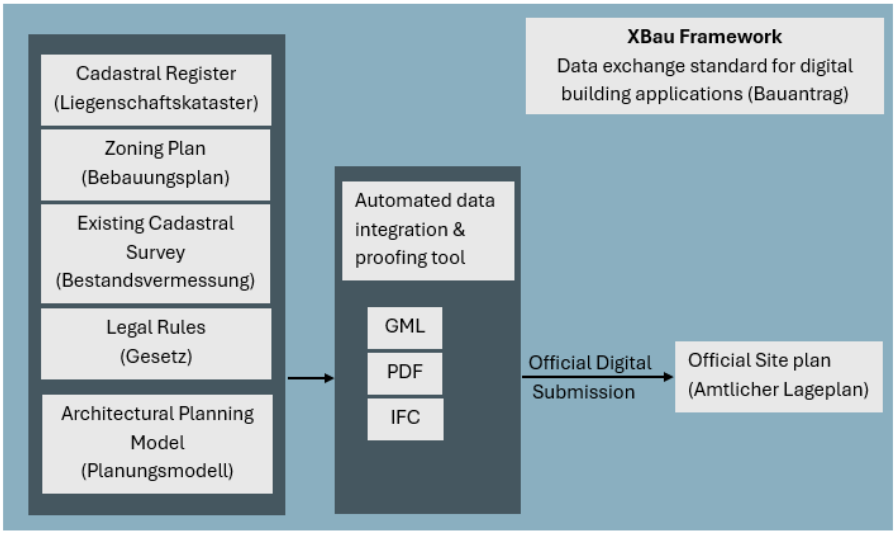


Figure 3: Processing pipeline for generating an official digital site plan.

This research focuses on the geometrical components of the application. Key elements from the site plan include: building lines, building boundary lines, property-line, and LOD2 (Level of Detail 2) CityGML building models with terrain.

These data are standardized through Germany's cadastral and planning systems. The AAA model (AFIS-ALKIS-ATKIS) provides legal and topographic baselines—especially ALKIS for parcels and boundaries. Planning data such as building lines are delivered via the national *XPlanung* standard, used for digital land-use and development plans (*Bebauungspläne*). LOD2 CityGML data is obtained from *Geobasis NRW*, the authoritative agency in North Rhine-Westphalia.

5.2 Geo-referencing of the IFC Model

To ensure compatibility with official datasets (e.g., ALKIS, CityGML), the IFC model is geo-referenced to German standard systems: ETRS89 / UTM Zone 33N (EPSG:25833) for horizontal and DHHN2016 for vertical reference. Using the `IfcMapConversion` class (IFC4+; see Section 4), the model's local coordinate reference system is mapped to the projected CRS through translation, rotation, and scaling. This enables spatial alignment with authoritative geodata while preserving the model's internal coordinate logic.

5.3 Schema Transformation between CityGML and IFC Data Models

1. Input Data Acquisition and Preprocessing

Official site information—parcel (GML coordinates), cadastral attributes (ALKIS IDs), and regulatory lines (building line, building boundary line) geometry are extracted using GML parsers, validated for topological correctness, and stored as intermediate geometric primitives like Cartesian points. See figure 4.

Data Mapping and Semantic Annotation

CityGML features are converted into IFC by mapping property-lines to `IfcSite` (using its `IfcSite.Representation`) encoding zoning attributes derived from XPlanung into `IfcPropertySetDefinition`, and representing regulatory constraints as surface geometry under `IfcProxyElement`. See figure 5. (Figure 4 and 5 are made with QGIS)

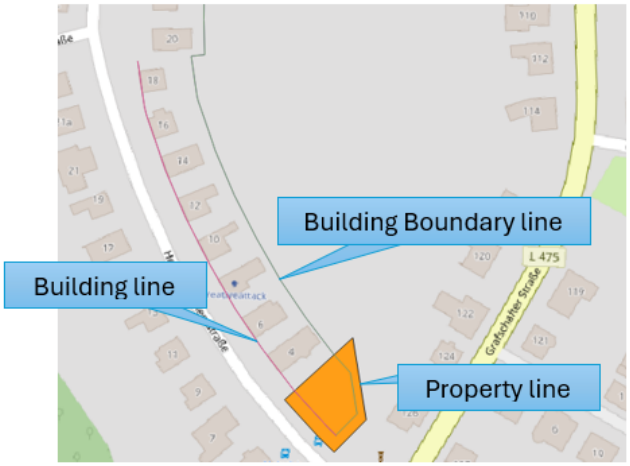


Figure 4: Site plan constraints in GIS data model

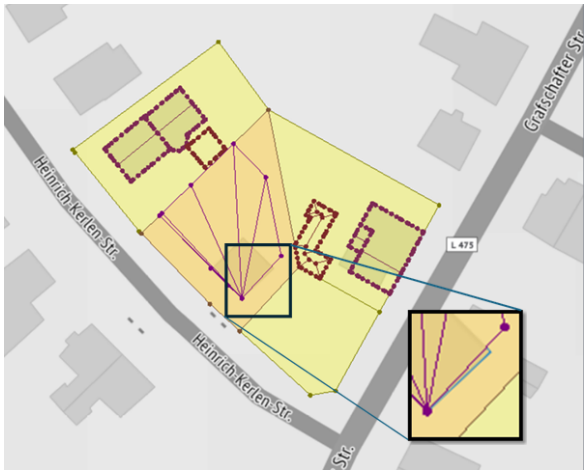


Figure 5: Parcels and buildings in IFC data-model.

3. Addressing Linear Constraints in IFC

Instance classes such as `IfcExtrudedAreaSolid` and `IfcSweptSolid` is challenging because they require closed planar profiles with thickness. Since linear constraints lack area or volume, they cannot be modeled directly without introducing artificial thickness, which in turn leads to approximation errors and a reduction in buildable volume.

4. Boundary Representation (BRep) Implementation

To accurately preserve regulatory constraints (Planzeichnungsverordnung § 3)[11], a pure Boundary Representation (BRep) approach is implemented using `IfcFacetedBrep`; see Algorithm 1 [12]. The final buildable volume, which serves as a spatial envelope for architectural design, is computed by applying constraints such as the building line, building boundary line and property-line boundaries. Additional parameters, including building height limits and setbacks from existing structures, can be incorporated in future developments of the model. A visual representation of the resulting buildable volume is shown in figure 6.

5.4 Validation of Models Against XPlanung Requirements

The resulting IFC model includes accurately defined B-Rep volumes that represent the permissible building envelopes. Automated scripts can traverse these volumes to extract dimensional parameters such as maximum allowed heights and buildable volumes and perform compliance checks based on spatial planning constraints derived from XPlanung data (see Section 5.1).

Geometric inconsistencies, such as elements extending beyond the defined regulatory B-Rep constraints, are flagged for revision. As shown in figures 7 and 8, the modeled geometry complies with

Algorithm 1 Boundary Representation in IFC for allowed buildable Volume

Geometric inputs: building line, building boundary line, and property-line
Output: IFC-compliant volumetric representation using `IfcFacetedBrep`
Profile Construction:
    Concatenate building line, building boundary line, and property-line
    Form a closed planar loop of 3D Cartesian points
Shell Topology Definition:
for each looped surface **do**
        Create `IfcFaceOuterBound` or `IfcFaceBound`
        Define the face as an `IfcFace` instance
        Apply geo-referencing using 2D Helmert transformation.
end for
Solid Assembly:
Aggregate all faces into an `IfcClosedShell`
Instantiate `IfcFacetedBrep` using the shell.
Integration:
Insert the `IfcFacetedBrep` into an `IfcProxyElement`
Link the proxy to the `IfcSite` using `IfcRelContainedInSpatialStructure`, as per IFC standard.

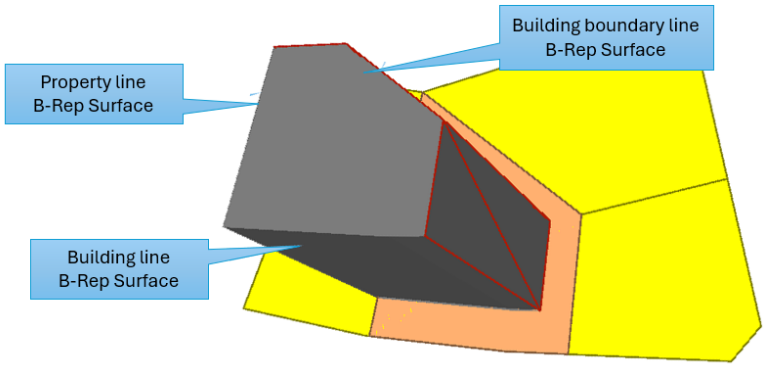


Figure 6: Buildable volume formed by regulatory constraints as surfaces, modeled into B-Rep solid.

the respective building line and property-line constraints. However, figures 5 and 9 indicate that the building boundary constraint has been violated.

Such validations can be automated using the `IfcOpenShell` library, which enables geometric analysis on IFC models. For advanced rule-based checking and regulatory validation, tools like Solibri Model Checker can be used to perform XPlanung-compliant spatial checks and generate comprehensive reports.

6 Results

The workflow converts CityGML-based parcel and zoning data into IFC proxy elements, modeling regulatory constraints as `IfcFacetedBrep` surfaces. This B-Rep approach preserves geometric accuracy and ensures alignment with legal envelopes. Geo-referencing via `IfcMapConversion` anchors the model in projected coordinates for integration with cadastral and topographic data. Automated

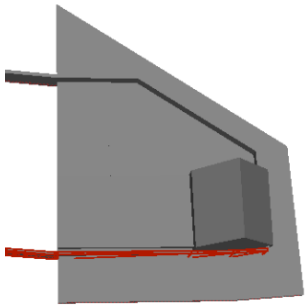


Figure 7: Building line

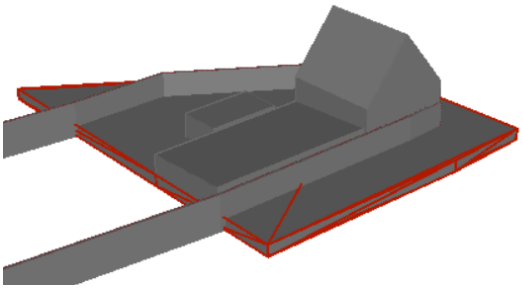


Figure 8: Property-line

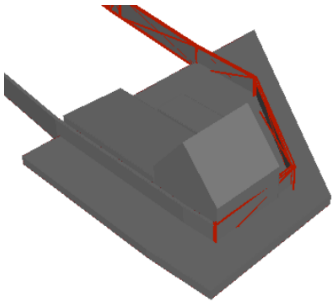


Figure 9: Building boundary line

checks using `IfcOpenShell` extract height and volume metrics, confirming compliance with building lines and parcel boundaries, while detecting violations such as building boundary line overreach. These results demonstrate the method's effectiveness for precise, rule-based validation within the IFC schema.

7 Conclusion and Outlook

The proposed method improves the accuracy and efficiency of digital permitting by enabling precise IFC modeling and automated constraint validation. It supports seamless integration of authoritative planning data (XPlanung) into BIM, and provides a XBau communication and data exchange standard ready 3D site plan.

However, several limitations remain. Different municipalities may interpret zoning regulations differently, which can complicate standardized rule checking. Current checks focus on basic constraints such as building lines, parcel boundaries, and height limits, while other rules required for building permits. Such as usage restrictions, protected or heritage areas, and public infrastructure constraints, are not yet fully integrated. In the scope of this research, Machine-readable interpretation of legal texts is still limited. Future work will implement the CityGML 3.0 model capabilities to extend validation to cover additional rules, such as height limits and setback requirements. Integration with commercial platforms like Solibri could enhance compliance reporting. This approach marks a step toward fully automated permitting, advancing Germany's digital transformation in construction.

Acknowledgements

The authors express their gratitude to our project partners: Bund der Öffentlich bestellten Vermessungsingenieure (BDVI), Vermessungsbüro KDS – Semler · Schult Öffentlich bestellte Vermessungsingenieure (ÖbVI) and the Technical University of Munich (TUM) for their contributions. This research was supported by funding from the German Federal Institute for Research on Building, Urban Affairs and Spatial Development (grant number 10.08.18.7-24.23.3).

References

[1] J. Fauth and L. Soibelman, "Conceptual framework for building permit process modeling: Lessons learned from a comparison between germany and the united states regarding the as-is building

permit processes", *Buildings*, vol. 12, no. 5, 2022. DOI: 10.3390/buildings12050638 [Online]. Available: https://www.mdpi.com/2075-5309/12/5/638

[2] Bundesinstitut für Bau-, Stadt- und Raumforschung, *Förderprogramm „Zukunft Bau" – Projekt 3D-Lageplan zum Baugesuch*, Projektsteckbrief, BBSR, 2024. [Online]. Available: https://www.3d-lageplan.org/zukunft-bau-forschungsprojekt

[3] M. Ataide, O. Braholli, and D. Siegele, "Digital transformation of building permits: Current status, maturity, and future prospects", *Buildings*, vol. 13, no. 10, 2023. DOI: 10.3390/buildings13102554 [Online]. Available: https://www.mdpi.com/2075-5309/13/10/2554

[4] Y. D. Altıntaş and M. E. İlal, "Integrating building and context information for automated zoning code checking: A review", *Journal of Information Technology in Construction (ITcon)*, vol. 27, pp. 548–570, 2022. DOI: 10.36680/j.itcon.2022.027 [Online]. Available: https://doi.org/10.36680/j.itcon.2022.027

[5] F. Noardo, T. Wu, K. A. Ohori, T. Krijnen, and J. Stoter, "Ifc models for semi-automating common planning checks for building permits", *Automation in Construction*, vol. 134, p. 104 097, 2022. DOI: 10.1016/j.autcon.2021.104097 [Online]. Available: https://doi.org/10.1016/j.autcon.2021.104097

[6] P. Azari, S. Li, A. Shaker, and S. Sattar, "Georeferencing building information models for bim/gis integration: A review of methods and tools", *ISPRS International Journal of Geo-Information*, vol. 14, no. 5, 2025. DOI: 10.3390/ijgi14050180 [Online]. Available: https://www.mdpi.com/2220-9964/14/5/180

[7] S. Comai, S. Mastrolembo Ventura, F. Noardo, K. Raja, and A. L. C. Ciribini, "Definition of bim and 3d city model information requirements for digital building permits", in *Proceedings of the Digital Building Permit Conference 2024*, Barcelona, 2024, pp. 193–199. [Online]. Available: https://iris.unibs.it/retrieve/8ef85646-a23b-4826-a14f-2491c4cb5cee/2024_Comai%20et%20al_DBP24.pdf

[8] H. Tauscher, N. E. Krieger, and F. Boochs, "A modular graph transformation rule set for ifc-to-citygml conversion", *Transactions in GIS*, vol. 25, no. 1, pp. 79–109, 2021. DOI: 10.1111/tgis.12723 [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1111/tgis.12723

[9] A. Rose, R. Becker, and A. Donaubauer, "3D-Lageplan zum Baugesuch", in *Leitfaden Geodäsie und BIM*, ser. DVW-Merkblatt 11-2021, R. Kaden et al., editors, Augsburg: Wißner Verlag, 2021, pp. 103–111. [Online]. Available: https://dvw.de/BIM-Leitfaden.pdf

[10] F. Müller, "Entwicklung einer citygml application domain extension: Der amtliche lageplan in 3d", Master's thesis, Technische Universität München, Munich, Germany, Jan. 2021.

[11] Bundesministerium des Innern, für Bau und Heimat. "Verordnung über die zur durchführung der bauleitplanung erforderlichen unterlagen (planzeichenverordnung – planzv)", Accessed: Jun. 25, 2025. [Online]. Available: https://www.gesetze-im-internet.de/planzv_90/BJNR000580991.html

[12] F. Ahmad, *Building permit application algorithm 1: Boundary representation in ifc*, https://github.com/tofatma/BuildingPermitapplication, 2025.

Prozessoptimierung in der BIM-basierten Tragwerksplanung: Entwicklung eines IFC-SAF-Konverters

Niklas Maczyewski[1], Dominik Schlütter[2], Jérôme Frisch[2] and Christoph van Treeck[2]
[1]RWTH Aachen University, Aachen, Germany
[2]Institute of Energy Efficiency and Sustainable Building (E3D), RWTH Aachen University, Aachen, Germany
E-mail(s): niklas.maczyewski@rwth-aachen.de

Abstract: Trotz der zunehmenden Adaption von openBIM in der Bauindustrie behindert ein latenter Medienbruch weiterhin die Effizienz tragwerksplanerischer Workflows. Vor der numerischen Berechnung ist es erforderlich, das architektonische IFC-Entwurfsmodell manuell in ein proprietäres 3D-Modell zu transformieren, um es anschließend in ein Analysemodell zu überführen. Die Bemessungsergebnisse sowie geometrische Veränderungen können sodann über geschlossene Schnittstellen in das proprietäre 3D-Modell zurückgeführt werden.

Der vorliegende Beitrag schließt diese Prozesslücke durch die Konzeption und prototypische Implementierung eines IFC-SAF-Konverters, der global eindeutige Identifikatoren (IFC-GUIDs) aus dem IFC-Entwurfsmodell in das erzeugte Structural Analysis Format (SAF)-Analysemodell überführt. Damit wird eine herstellerunabhängige Zuordnung zwischen den heterogenen Modellen gewährleistet.

Die erarbeitete Prozesslogik invertiert den konventionellen Ablauf in der BIM-basierten Tragwerksplanung. Das proprietäre 3D-Modell muss nicht mehr vorab manuell erstellt, sondern kann automatisiert aus dem Analysemodell generiert werden. Dadurch entfallen manuelle Vormodellierungen, Iterationszeiten werden signifikant verkürzt und die Interoperabilität zwischen CAD- und CAE-Umgebungen verschiedener nicht miteinander kooperierender Softwarehäuser wird substantiell verbessert.

Keywords: structural analysis, bim, processes

DOI: 10.18154/RWTH-CONV-254926. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Einleitung

Digitale Tragwerksanalysen gehören zu den am häufigsten genutzten BIM-Anwendungen in Planungsprozessen [1]. Dennoch zeigen Querschnittsdaten aus 2019, dass lediglich 38% der befragten Ingenieurbüros BIM konsequent in ihren Planungsprozessen einsetzen [2].

Diese Diskrepanz macht deutlich, dass das umfassende Potenzial der 3D-Planung innerhalb der Strukturanalyse noch nicht ausgeschöpft wird. Anstatt sich im BIM-basierten Tragwerksentwurf auf die

Optimierung der dreidimensionalen Berechnung zu konzentrieren, sollten die Möglichkeiten der 3D-Planung genutzt werden, um bestehende Prozessketten im Zusammenhang mit der Strukturanalyse zu verbessern. Die Ableitung eines Analysemodells gehört dabei grundsätzlich zu den Kernaufgaben der Tragwerksplanung im BIM-Prozess.

In den letzten Jahren wurden zahlreiche Anstrengungen unternommen, um durchgängige Infrastrukturen, Prozesse und Werkzeuge zu entwickeln, mit denen sich Informationen effizient aus 3D-Gebäudemodellen extrahieren und in CAE-Software verwenden lassen, doch die Umwandlung eines realitätsnahen architektonischen Entwurfsmodells in ein nutzbares Analysemodell bleibt häufig mühsam und zeitaufwändig.

Gerade im Bereich indirekter Verknüpfungen (openBIM) bestehen noch erhebliche Defizite, da kommerzielle Autorenwerkzeuge und CAE-Anwendungen keine robusten, plattformübergreifenden Identifikatoren bereitstellen, sodass Änderungen, die sich durch die Berechnung und Bemessung am Analysemodell ergeben, nicht ohne umfangreichen manuellen Aufwand in das 3D-Modell zurückgeführt werden können.

Die Auslagerung der Verknüpfung in standardisierte Austauschdateien ermöglicht eine herstellerunabhängige Nutzung unterschiedlicher CAD- und CAE-Anwendungen. Somit entfällt die Notwendigkeit, proprietäre Softwaresuiten desselben oder kooperierender Hersteller zu nutzen. Zugleich wird die gewerkeübergreifende Weiterverwendung von Planungsdaten gefördert.

Vor diesem Hintergrund wurde, basierend auf der kombinierten Nutzung der Engine von That Open Company und der Open-Source Bibliothek IfcOpenShell ein Konverter entwickelt, der konsequent offene Austauschformate verwendet. Dieser ermöglicht die Ableitung belastbarer Analysemodelle aus physischen 3D-Entwurfsmodellen, stellt die Verknüpfung der heterogenen Modelltypen innerhalb der offenen Austauschdateien sicher und schafft damit die Grundlage für einen optimierten Planungsablauf in der BIM-basierten Tragwerksplanung, indem vorhandene Modellinformationen nicht erneut generiert, sondern effizient weiter genutzt werden können.

2 Technischer Hintergrund

2.1 Prozesse in der BIM-basierten Tragwerksplanung

Die Tragwerksplanung ist ein integraler Bestandteil des Bauprozesses und verantwortet maßgeblich die strukturelle Standsicherheit, Gebrauchstauglichkeit, Dauerhaftigkeit sowie Wirtschaftlichkeit eines Rohbauwerks [3]. In diesem Zusammenhang wird ein architektonisches Entwurfskonzept in ein Tragwerk überführt bzw. werden Tragstrukturen aus Bestandsobjekten abgeleitet und bemessen [3], [4]. Mit der Nutzung eines dreidimensionalen, digitalen Gebäudemodells im Rahmen des BIM-Planungsprozesses verlagert sich die Aufgabe des Tragwerksplaners von der manuellen Berechnung hin zur Modellbildung und -validierung [5].

Die wesentlichen Arbeitsschritte gliedern sich dabei in drei Abschnitte:

- die Erstellung eines proprietären 3D-Modells,
- die Ableitung eines Analysemodells und
- die Umsetzung sowie Bemessung in Berechnungsmodelle [6].

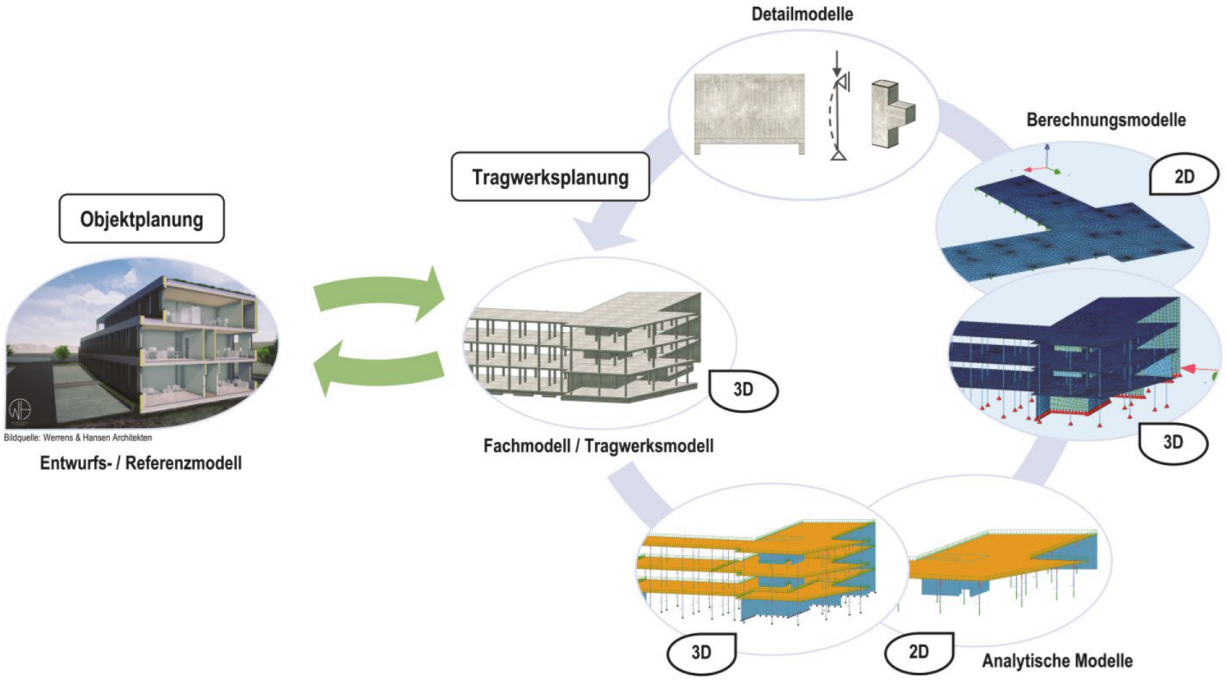


Abbildung 1: Digitale Modelle in der Objekt- und Tragwerksplanung [7].

Im physischen 3D-Modell werden statisch relevante Bauteile als Volumenobjekte abgebildet (vgl. Abbildung 1) [8]. Das Modell wird auf Grundlage der Planungsunterlagen des Entwurfs erstellt und fortlaufend um tragwerksspezifische Attribute (z. B. Materialien, Feuerwiderstandsklassen, Angaben zur Dauerhaftigkeit, Expositionsklassen, ggf. Anschluss- und Auflagerbedingungen, Querschnitte etc.) ergänzt.

Zur adäquaten Analyse des Tragverhaltens, Bestimmung des Spannungs-Dehnungs-Zustandes sowie Lösung von Entwurfs- oder Prüfaufgaben benötigt der Tragwerksplaner jedoch eine angemessene mathematische Abstraktion der tatsächlichen Struktur.

Das Analysemodell stellt eine abstrahierte Repräsentation des physischen 3D-Modells dar, in der die volumetrische Geometrie gezielt auf eindimensionale Stab- bzw. zweidimensionale Platten- und Schalenelemente reduziert wird, um die Tragstruktur rechentechnisch effizient abzubilden (vgl. Abbildung 1) [7], [9].

Zur Ermittlung der maßgebenden Spannungs-, Deformations- und Stabilitätsgrößen sowie zur Führung der entsprechenden Nachweise wird das Analysemodell unter Berücksichtigung der gewählten Berechnungs- und Bemessungsverfahren innerhalb der CAE-Software in Gesamt- oder Teilberechnungsmodelle überführt und bemessen [7].

2.2 Herausforderungen bei der Ableitung von Analysemodellen aus physischen 3D-Modellen

Eine der zentralen Problematiken bei der automatischen Transformation vom physischen 3D- zum Analysemodell besteht in der Bildung eines zusammenhängenden logischen Systems. In physischen Fachmodellen werden Strukturbauteile schalkantenbezogen abgebildet [8]. Dies verursacht bei der Ableitung eines Analysemodells eine Reihe von Herausforderungen:

- Systemlinienproblematik: Differenzen zwischen den analytischen Knoten benachbarter Bauteile infolge der dimensionsreduzierten Darstellung.
- Fehlen statisch relevanter Informationen: Entwurfsmodelle besitzen in der Regel nicht alle benötigten Informationen für die statische Bemessung (z. B. Materialien und Querschnitten, ggf. Angaben zu den Anschluss- und Auflagerbedingungen etc.).
- Modellkontrolle: Ungenauigkeiten bei der Modellierung von physischen 3D-Modellen können dazu führen, dass Knoten im Analysemodell entstehen, die zwar sehr nahe beieinander liegen, aber trotzdem nicht richtig miteinander verbunden sind.
- Modellverknüpfung: Bi-direktionale Verknüpfung von 3D- und Analysemodell trotz domainspezifischer Divergenzen.

2.3 Structural Analysis Format

Das IFC-Schema hat sich als Standard für den offenen Austausch von Daten in der Bau- und Immobilienbranche etabliert. Eine gewerkeübergreifende Integration aller domänenspezifischen Modelle in ein einziges Datenmodell würde jedoch die inzwischen ohnehin hohe Komplexität und den Umfang des IFC-Standards erheblich steigern und somit dessen Praxistauglichkeit einschränken. Dies betrifft insbesondere Modelle, die auf einer abstrakteren Darstellungsebene operieren als die physische Geometrie eines 3D-Bauwerksmodells. Vor diesem Hintergrund erkannte die Nemetschek Group, dass ein separates, offenes Austauschformat für Analysemodelle erforderlich ist, um den Anforderungen der Tragwerksplanung in einem BIM-Kontext gerecht zu werden. Im Jahr 2017 wurde auf Grund dessen unter der Koordination von SCIA (Teil der Nemetschek Group) das Structural Analysis Format entwickelt, welches seither kontinuierlich weiterentwickelt wird. Bei einer SAF-Datei handelt es sich um eine standardisierte Excel-Tabelle in der Form einer relationalen Datenbank, die eine maschinen- und menschenlesbare Abbildung strukturanalytisch relevanter Daten ermöglicht. Analog zum IFC-Datenmodell definiert das SAF spezifische Relationen (Klassen) für analytische und nicht-physische Objekte [10].

Jede Relation besitzt eine fest definierte Menge von Attributen. Innerhalb der Excel-Arbeitsmappen gilt (vgl. Abbildung 2):

- Spalte – repräsentiert ein Attribut der Relation
- Zeile – repräsentiert einen Tupel (Objekt) der Relation
- Zelle – repräsentiert einen spezifischen Wert eines Attributs

Im Sinne einer relationalen Datenbank werden verschiedene Arbeitsmappen dabei unter Verwendung von Fremd- und Primärschlüsseln logisch miteinander verknüpft.

3 Methodische Umsetzung

Die Implementierung des IFC-SAF-Konverters erfolgte in einem testmodellgetriebenen Entwicklungsprozess. Dazu wurden zunächst unter Berücksichtigung der Herausforderungen funktionale und datenbezogene Anforderungen für die Ableitung von Analysemodellen aus der Entwurfsgeometrie

Zeile = Tupel Spalte = Attribut Zelle = Attributwert

	A	B	C	D	E	F	G	H
1	Name	Type	Subtype	Quality	Unit mass [kg/m3]	E modulus [MPa]	G modulus [MPa]	Poisson coefficient
2	C20/25	Concrete		C20/25	2500,0	31500,000000	13125,000000	0,200
3	C20/30	Concrete		C25/30	2500,0	31500,000000	13125,000000	0,200
4	C30/37	Concrete		C30/37	2500,0	31500,000000	13125,000000	0,200
5	S235	Steel		S235	7850,0	210000,000000	80769,230769	0,300
6	S355	Steel		S355	7850,0	210000,000000	80769,230769	0,300
7	EN-AW 5083	Aluminium		EN-AW 5083	2700,0	70000,000000	26923,076923	0,300
8	EN-AW 8011A	Aluminium		EN-AW 8011A	2700,0	70000,000000	26923,076923	0,300
9	D30 (EN 338)	Timber		D30 (EN 338)	420,0	11000,000000	690,000000	0,000
10	C22 (EN 338)	Timber		C22 (EN 338)	420,0	11000,000000	690,000000	0,000
11	GL 30c (EN 14080)	Timber		GL 30c (EN 14080)	420,0	11000,000000	690,000000	0,000
12	Masonry	Masonry		Masonry	1600,0	3100,000000	1240,000000	0,250

Abbildung 2: Aufbau einer Arbeitsmappe am Beispiel der Relation StructuralMaterial

sowie für die Aufrechterhaltung bidirektionaler Verknüpfungen zwischen beiden Modelltypen definiert. Den Ausgangspunkt bilden auf Grund der hohen Verbreitung Entwurfsmodelle im IFC-2x3-Schema (Coordination View 2.0), dessen IFC-Entitäten semantisch auf die korrespondierenden Relationen des SAF abgebildet werden (vgl. Abbildung 3).

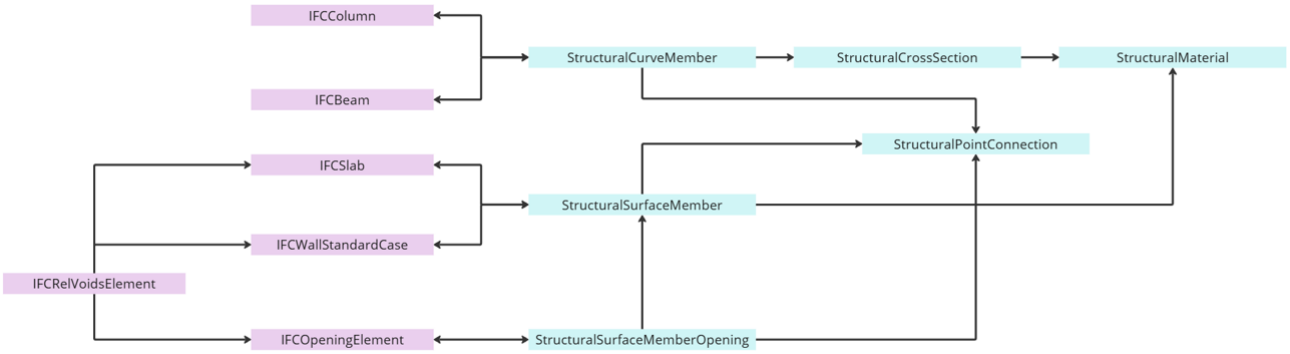


Abbildung 3: Beziehungsdarstellungen und Zuordnungen von IFC-Klassen und SAF-Relationen

Auf Grundlage der Standardisierungen des IFC- und SAF-Schemas wurden die erarbeiteten Anforderungen unter Verwendung von Open-Source-Technologien in einen Konverter überführt. Die Implementierung erfolgte dabei anhand eines Testmodells. Die generierten SAF-Analysemodelle wurden mit Viewern sowie kompatibler CAE-Software validiert. Im Anschluss wurde ein Skript entwickelt, das die Informationen aus den erzeugten SAF-Dateien ausliest und daraus proprietäre 3D-Modelle generiert. Die erarbeitete Prozesslogik wurde in einem End-to-End-Test abschließend geprüft.

4 Implementierung des IFC-SAF-Konverters

Die Implementierung überführt die in der methodischen Umsetzung beschriebenen Spezifikationen in eine lauffähige Systemarchitektur. Ein webbasiertes Frontend (Converter.ts) lädt dabei IFC-Dateien, visualisiert sie im Viewer und stellt Funktionen für das Entfernen statisch unrelevanter Objekte (Cleaning) sowie die Eingabe fehlender statisch relevanter Attribute (Materials.ts, Crosssections.ts) bereit. Alle Eingaben werden gebündelt und per REST an eine leichte API (flaskserver.py) übertragen. Diese startet die Konvertierungs-Engine (SAFTemplate.py), welche die bereitgestellten Modellinformationen parst, relevante Informationen extrahiert und diese strukturiert zusammen mit den Eingaben

unter Berücksichtigung der SAF-Standardisierungen in eine Excel-Datei schreibt. Der operative Datenfluss umfasst das Laden und Prüfen der Eingabedatei, das optional interaktive Nachführen von Attributen, die Übergabe an die Serverkomponenten, die Befüllung der Zielformate und die Rückgabe eines Download-Hinweises an das Frontend. Die umgesetzte Systemarchitektur ist in Abbildung 4 dargestellt.

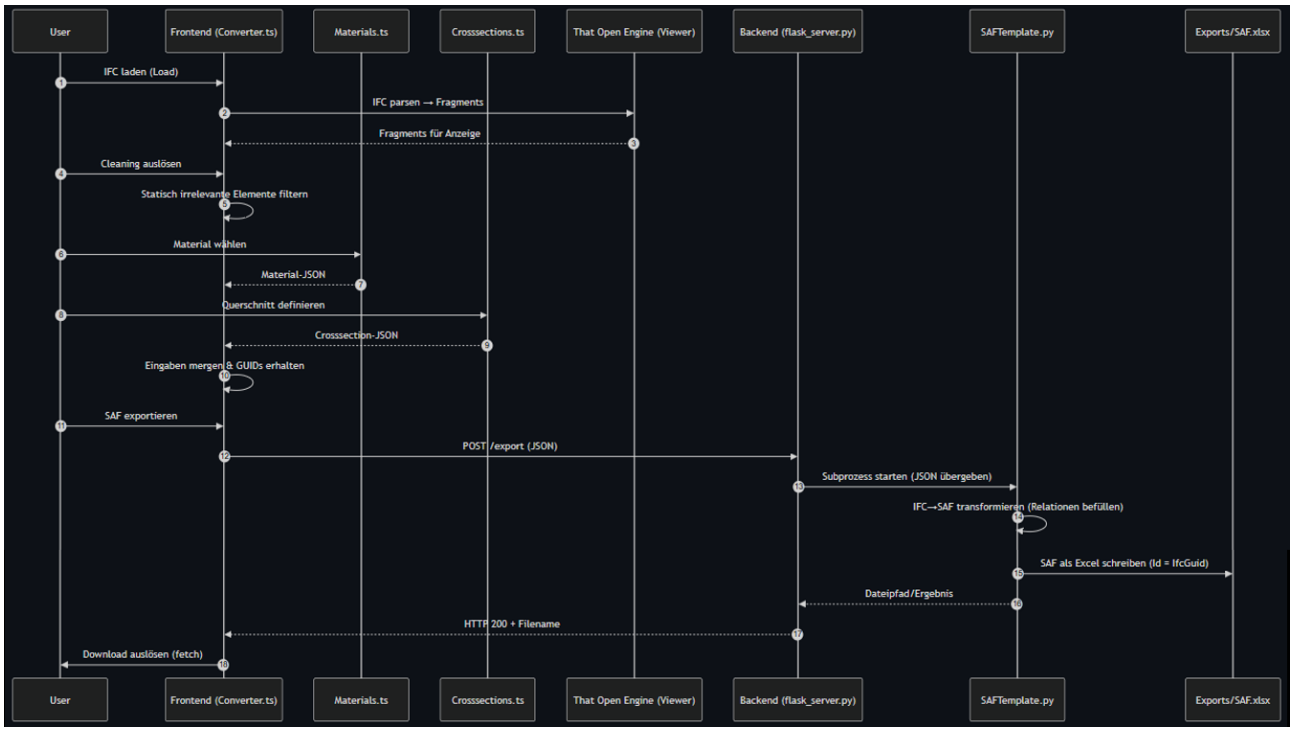


Abbildung 4: Softwarearchitektur des IFC-SAF-Konverters

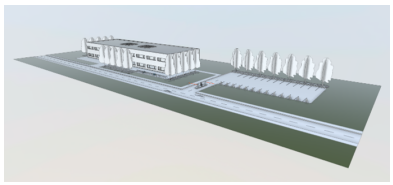
5 Gegenüberstellung konventioneller und optimierter openBIM-Prozess in der Tragwerksplanung

Im konventionellen openBIM-Ablauf erhält der Tragwerksplaner zunächst das IFC-Entwurfsmodell des Architekten, modelliert daraus manuell ein proprietäres 3D-Modell in seiner CAD-Software und erstellt anschließend manuell oder durch die Verwendung von Tools das Analysemodell in der CAE-Umgebung. Geometrische Änderungen, die sich aus der Berechnung ergeben, müssen wieder manuell oder durch die Verwendung geschlossener Verbindungen in das 3D-Modell übertragen werden, sodass sich der Modellierungsaufwand mit jeder Iteration erhöht; zugleich verlässt der Ingenieur damit seinen eigentlichen Kernauftrag, nämlich die statische Analyse.

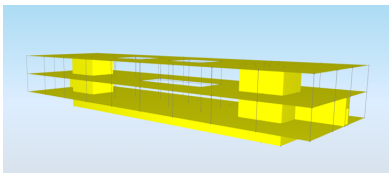
Der optimierte Prozess verschiebt diese Reihenfolge: Ausgehend vom gleichen IFC-Entwurfsmodell wird mithilfe des entwickelten IFC-SAF-Konverters sofort ein Analysemodell abgeleitet. Die dabei erzeugte SAF-Datei übernimmt die ursprünglichen IFC-GUIDs, sodass Entwurf-, Analyse- und Berechnungsmodell digital verknüpft bleiben. Nach der ersten Berechnung wird das aktualisierte Analysemodell erneut als SAF exportiert; aus dieser Datei kann innerhalb der CAD-Software das verknüpfte proprietäre 3D-Modell über einen skriptbasierten Ansatz nun automatisch generiert werden, da sich alle benötigten Informationen in der SAF-Datei befinden (vgl. Abbildung 5). Künftige Änderungen

können dank der GUID-Kette ohne manuelle Nacharbeit zurück in das 3D-Modell gespielt werden, wodurch redundante Modellierungen entfallen und statische Probleme bereits vor der detaillierten Strukturmodellierung erkannt und an den Architekten kommuniziert werden können.

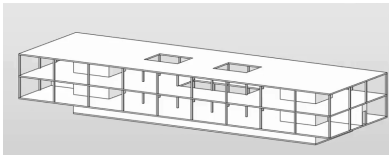
Damit ersetzt der Konverter doppelte Modellierungsarbeiten durch die Weiterverwendung vorhandener Informationen und hat das Potenzial, Iterationszyklen signifikant zu verkürzen.



(a) IFC-Entwurf



(b) generiertes Analysemodell



(c) proprietäres 3D-Modell

Abbildung 5: Vom IFC-Entwurf über das generierte Analysemodell zum proprietären 3D-Modell.

6 Zusammenfassung und Ausblick

Trotz der zunehmenden Akzeptanz von openBIM-Workflows beruhen die heute gebräuchlichen Kopplungen zwischen statischem Analyse- und physischem 3D-Modell auf proprietären Mechanismen. Ein automatisches Ableiten eines verknüpften Analysemodells ist daher nur möglich, wenn zuvor ein manuell erstelltes proprietäres 3D-Modell vorliegt.

Der vorgestellte Ansatz durchbricht diesen methodologischen Rahmen: Ausgehend vom IFC-Entwurfsmodell wird ein SAF-Analysemodell generiert, dessen Verknüpfungen durch die Übernahme der IFC-GUIDs in die Austauschdateien eingebettet sind. Hierzu wurde ein Konverter entwickelt, der ausschließlich auf offenen Datenformaten (IFC 2x3 CV 2.0, SAF) und frei verfügbaren Technologien aufsetzt. Er ist in einen optimierten Workflow eingebettet, der vorhandene Modellinformationen konsequent weiterverwendet.

Gegenüber dem konventionellen openBIM-Workflow in der Tragwerksplanung wird der manuelle Modellierungsaufwand signifikant reduziert, iterative Änderungszyklen werden verkürzt und statische Inkonsistenzen können deutlich früher erkannt werden.

Nichtsdestotrotz erfordert die nachhaltige Implementierung des Ansatzes eine weitergehende Standardisierung, um domänenbedingte Modelldivergenzen zwischen Analyse- und 3D-Modell systematisch abzufangen. Komplexere geometrische Unterschiede, wie zum Beispiel Versätze, können meist nicht allein durch die standardisierten SAF-Attribute abgebildet werden. Um diese Informationen verlustfrei zu übertragen und bei der Modellierung bzw. Aktualisierung entsprechend zu berücksichtigen, wären über die derzeitigen Konventionen im SAF hinausgehende Informationen nötig. Darüber hinaus muss grundsätzlich für diesen Ansatz gewährleistet sein, dass im Rahmen von Im- und Exportvorgängen innerhalb der CAE-Anwendungen eine durchgängig stabile Zuordnung der SAF-Relationen besteht, sodass geometrische Modellinformationen unverzerrt in die zugrundeliegenden Modelle überführt werden können.

Acknowledgements

Teile dieser Arbeit wurden im Rahmen einer Masterarbeit am Lehrstuhl für Energieeffizientes Bauen E3D der RWTH Aachen entwickelt.

Literatur

[1] I. J. Ramaji und A. M. Memari, »Interpretation of structural analytical models from the coordination view in building information models«, *Automation in Construction*, Jg. 90, S. 117–133, 2018. DOI: 10.1016/j.autcon.2018.02.025 Adresse: https://doi.org/10.1016/j.autcon.2018.02.025

[2] F. Bialas, V. Wapelhorst, S. Brokbals und I. Čadež, »Quantitative Querschnittsstudie zur BIM-Anwendung in Planungsbüros«, *Bautechnik*, Jg. 96, Nr. 3, S. 229–238, 2019. DOI: 10.1002/bate.201800103 Adresse: https://onlinelibrary.wiley.com/doi/10.1002/bate.201800103

[3] G. Jahnke, *Grundlagen der Tragwerksplanung: Von der Entwurfsidee über das Tragsystem zum Tragwerk* (erfolgreich studieren). Wiesbaden: Springer Vieweg, 2023. DOI: 10.1007/978-3-658-38165-3 Adresse: https://doi.org/10.1007/978-3-658-38165-3

[4] A. Bögle und K.-E. Kurrer, »Das strukturale Komponieren von Tragwerken bei Jörg Schlaich«, *Beton- und Stahlbetonbau*, Jg. 109, Nr. 11, S. 829–837, 2014. DOI: 10.1002/best.201400089 Adresse: https://doi.org/10.1002/best.201400089

[5] M. R. Bischoff, »Statik am Gesamtmodell: Modellierung, Berechnung und Kontrolle«, *Der Prüfingenieur*, Jg. 36, S. 27–34, Apr. 2010. Adresse: https://bvpi.de/bvpi/downloads/der-pruefingenieur/Pruefingenieur_36.pdf

[6] R. Kepplin, M. Schnellenbach-Held und M. Held, »Building Information Modeling – Umsetzung in der Tragwerksplanung«, *Bautechnik*, Jg. 94, Nr. 4, S. 220–226, Apr. 2017. DOI: 10.1002/bate.201700003 Adresse: https://doi.org/10.1002/bate.201700003

[7] J. Stracke und R. Kepplin, »Der BIM-Prozess in der Tragwerksplanung«, *Beton- und Stahlbetonbau*, Jg. 115, Nr. 4, S. 324–331, Apr. 2020. DOI: 10.1002/best.201900097 Adresse: https://doi.org/10.1002/best.201900097

[8] A. Borrmann, M. König, C. Koch und J. Beetz, Hrsg., *Building Information Modeling: Technologische Grundlagen und industrielle Praxis* (VDI-Buch), 2. Aufl. Wiesbaden: Springer Vieweg, 2021. DOI: 10.1007/978-3-658-33361-4 Adresse: https://doi.org/10.1007/978-3-658-33361-4

[9] C. van Treeck, R. Elixmann, K. Rudat, S. Hiller, S. Herkel und M. Berger, Hrsg., *Gebäude. Technik. Digital: Building Information Modeling* (VDI-Buch), 1. Aufl. Berlin Heidelberg: Springer Vieweg, 2016. DOI: 10.1007/978-3-662-52825-9 Adresse: https://link.springer.com/book/10.1007/978-3-662-52825-9

[10] SCIA. »SAF Documentation: Version 2.2.0«, besucht am 25. Jan. 2025. Adresse: https://www.saf.guide/en/stable/index.html

An API-Driven System for Real-Time Safety Monitoring on Construction Sites: Integrating Machine Learning and Industry Foundation Classes

Mukta Nisal[1] and Xiangyu Wang[1]
[1]Institut für Bauinformatik , TU Dresden University, Dresden, Germany
E-mail(s): mukta_rajan.nisal@tu-dresden.de, xiangyu.wang@tu.dresden.de

Abstract: This research focuses on improving construction site safety by developing an innovative real-time hazard detection system using machine learning models. At its core is a camera-based backend API that deploys machine learning models for object detection in an SQL database. When potential safety issues arise, the system triggers alerts and logs key data points such as timestamps and alert types. Alerts are grouped into short-term (24-hour) and long-term (1-month) categories, which offer actionable insights for immediate and ongoing safety improvements. By integrating advanced object detection methods like Yolov8 with efficient database management and API-driven functionality, the system offers an effective solution to monitor construction site safety in real time. Its ability to dynamically manage machine learning models and alert categorization functionality make it an essential tool for fostering safer workplaces. Future developments include integrating location data using compatible sensors to further enhance risk monitoring. Additionally, the paper discusses future integration with the Industry Foundation Classes, which could expand the system's capabilities by linking hazard detection to specific building elements and processes.

Keywords: construction site safety, real-time object detection, machine learning models, alert generation, SQL database

DOI: 10.18154/RWTH-CONV-254924. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

The construction industry remains one of the most hazardous sectors globally, with frequent safety lapses attributed to delayed hazard recognition, non-compliance, and the ever-changing site environment [1], [2]. Traditional safety approaches which are largely reactive and manual often do not provide the real-time awareness required to prevent accidents [1]. Recent advances in artificial intelligence (AI) and computer vision have opened new avenues for proactive safety monitoring. AI-driven object detection models, such as those in the YOLO (You Only Look Once) family, enables high-speed, accurate recognition of safety-critical elements like PPE compliance, unguarded edges and equipment proximity

[3], [4], [5].This study proposes a deployable, real-time construction safety detection framework using YOLOv8n integrated with a FastAPI backend. The system processes both helmet-mounted and static camera inputs, offering a robust detection architecture. It supports modular updates, adapts to site changes and provides actionable insights through dashboard integration [6]. The goal is to validate this system under varied site conditions and demonstrate its effectiveness in enhancing safety through immediate risk detection and response.

2 Literature Review

2.1 AI for Construction Safety Monitoring

Recent advances in artificial intelligence (AI) and computer vision have significantly reshaped safety monitoring in the high-risk construction environment. Alateeq et al. [6] utilized deep learning to identify site hazards with high precision, demonstrating scalable detection of personal protective equipment (PPE) violations. Similarly, Zhou et al. [2] conducted a comprehensive review of safety management strategies, supporting the integration of AI for proactive safety enforcement. Arabshahi et al.[7] emphasized the increasing adoption of sensing technologies in the construction industry, with sensors and AI forming the backbone of modern safety infrastructures. Boje et al. [8] advocated for semantic digital twins to facilitate real-time risk analysis and intervention. Karlapudi et al. [9] demonstrated how Internet of Things (IoT) devices, digital twins, and data fusion can be leveraged to improve safety on construction sites. This approach strongly supports the integration of AI-based hazard detection with Building Information Modeling (BIM), highlighting the need for interoperable systems that combine real-time monitoring with digital representations of construction processes.

2.2 Detection Models: CNNs and YOLO-based Approaches

Detection models such as convolutional neural networks (CNNs) and the YOLO (You Only Look Once) series are crucial in automating risk identification. Redmon et al. [3] pioneered YOLO, which delivers real-time object detection with minimal latency. Jocher [4] introduced YOLOv8, offering Open Neural Network Exchange (ONNX) compatibility and a modular design suitable for deployment on edge devices. Sohan et al. [10] reviewed the improvements of YOLOv8, particularly for complex construction environments. Talib et al. [11] proposed YOLOv8-CAB, achieving enhanced detection precision in dynamic scenes. Zhang et al. [5] demonstrated high-accuracy PPE detection using sensor-integrated CNNs. Parkash et al. [12] supported RFID-based systems for location-aware risk evaluation. Arabshahi et al. [7] also discussed wearable sensor systems that enable real-time proximity detection of hazards using Ultra-Wideband (UWB) and other localization technologies.

2.3 Alert Systems and Smart Integration

Arabshahi et al. [7] and Zhang et al. [5] demonstrated the integration of sensor data with AI-based real-time alert systems to reduce incident response times in construction settings. Talib et al. [11] further improved this concept with their YOLOv8 framework, which combines high-speed detection with enhanced alerting mechanisms suitable for dynamic site conditions. Despite progress, challenges remain: sensor noise, model robustness, and audit transparency. As Alateeq et al. [6] suggest, improved data fidelity and edge-device compatibility are essential for next-generation deployments.

2.4 Categorization of Hazards

Hossen [13] analyzed and categorized hazard situations and locations at construction sites in his master's thesis at TU Dresden. His work emphasizes that hazards can be systematically classified into categories such as physical, environmental, and procedural risks. This categorization approach highlights the importance of not only identifying hazards but also contextualizing them within their specific risk domains, thereby improving targeted mitigation strategies.

3 Methodology

This section outlines the technical foundation of our real-time hazard detection system tailored for high-risk construction sites. The pipeline integrates AI-driven object detection and a FastAPI microservice for scalable deployment.

3.1 FastAPI-Based Object Detection Service

The Hazard Detection API is a FastAPI-based microservice that provides real-time computer vision capabilities. It includes intuitive endpoints like `POST /detect` (submit images), `GET /yolo/models` (model metadata), and `POST /update_model` (YOLOv8 weight hot-swapping). The service fuses UWB tag coordinates with 360-degree helmet-mounted and static camera feeds, returning annotated bounding boxes, confidence scores, and 3D hazard positions. Analytical endpoints such as `/detection/alerts` and `/detection/analyze` allow retrieval of historical detection logs and safety summaries.

3.2 Object Detection Workflow

Figure 2 illustrates the end-to-end pipeline for object detection. It begins with data collection using 360-degree and standard cameras. The data collected are enriched through Open Experience tools like BauMangel and BauPhotos. Data was collected using two major tools:

1. **360-Degree Camera:** The foundation of object detection is the data collection process. This process was done with the help of two major tools: the 360-degree Camera and a standard camera. The 360-degree camera system from Open Experience GmbH is a panoramic device seamlessly integrated into a construction helmet. Equipped with two 12MP cameras and high-quality wide-angle lenses, this system ensures exceptional detail and clarity in every 360-degree image.

2. **Standard Camera:** In addition to the 360° camera, a standard camera provides object-focused views with varying fields of view, improving recognition under diverse conditions. Combined with panoramic data, this creates a rich dataset covering both contextual and detailed perspectives.

Figure 1: Helmet-mounted 360-degree camera for data acquisition

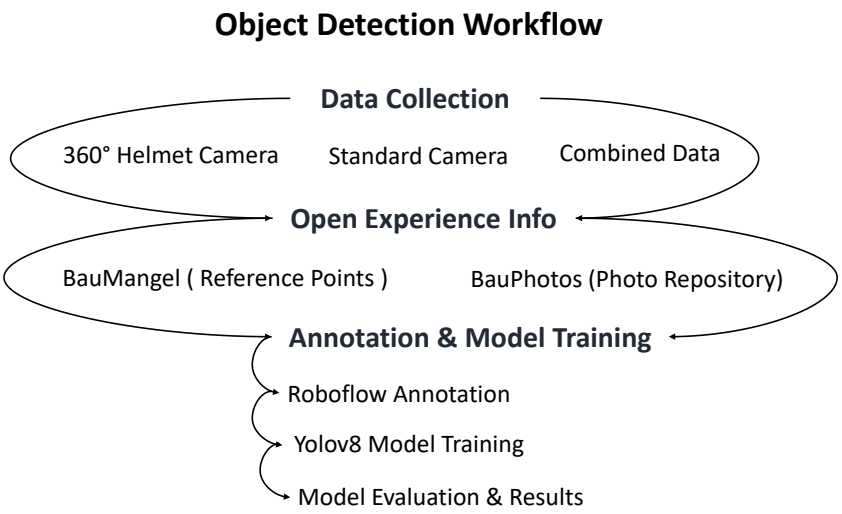


Figure 2: Object detection workflow pipeline from data collection to model evaluation.

3.3 Annotation and Model Training

1. **Roboflow Annotation:** (Figure 3) All collected images were uploaded to the RoboFlow Platform for manual annotation using web interface and bounding boxes were drawn around the critical safety classes such as `hazardous_object`, `big_obstacle`, `open_wires`. The final RoboFlow-annotated dataset was exported into a YOLOv8-compatible format.

Figure 3: Roboflow annotation interface showing bounding boxes

2. **YOLOv8 Model Training:** The annotated dataset was exported directly in YOLOv8 format, eliminating manual conversion steps. We selected the YOLOv8n (“nano”) variant for its balance of real-time inference speed and detection accuracy, leveraging its anchor-free heads and CSP-based feature fusion. Training was carried out via the Ultralytics YOLOv8 Python library, with hyperparameters tuned (learning rate, batch size, augmentation mix) over 300 epochs. Key performance metrics mAP@0.5 (0.75), precision (0.82), recall (0.78), and average inference time (50 ms) were monitored to guide further augmentation and learning-rate adjustments.

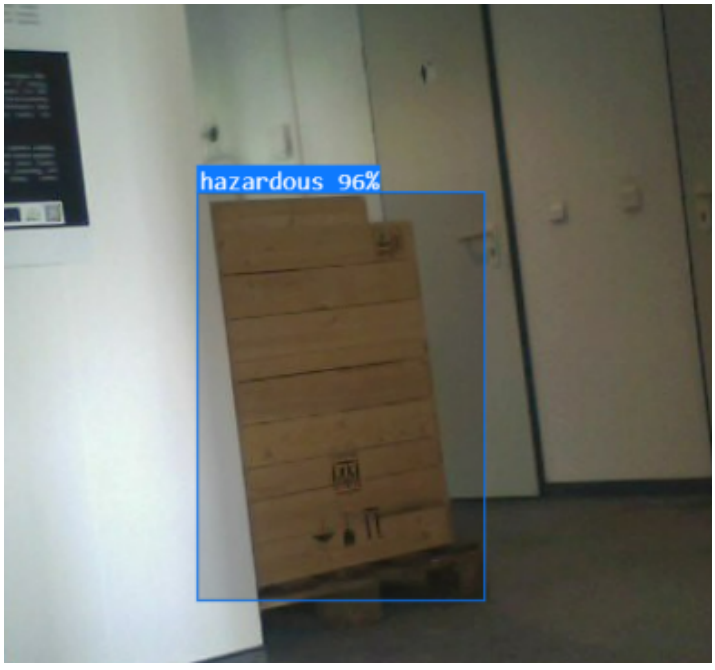


Figure 4: YOLOv8 model output on construction dataset

3.4 API Architecture Overview

The architecture, shown in Figure 5, illustrates the proposed hazard detection system comprising multiple interacting components, each assigned specific responsibilities to facilitate real-time safety monitoring and system responsiveness.

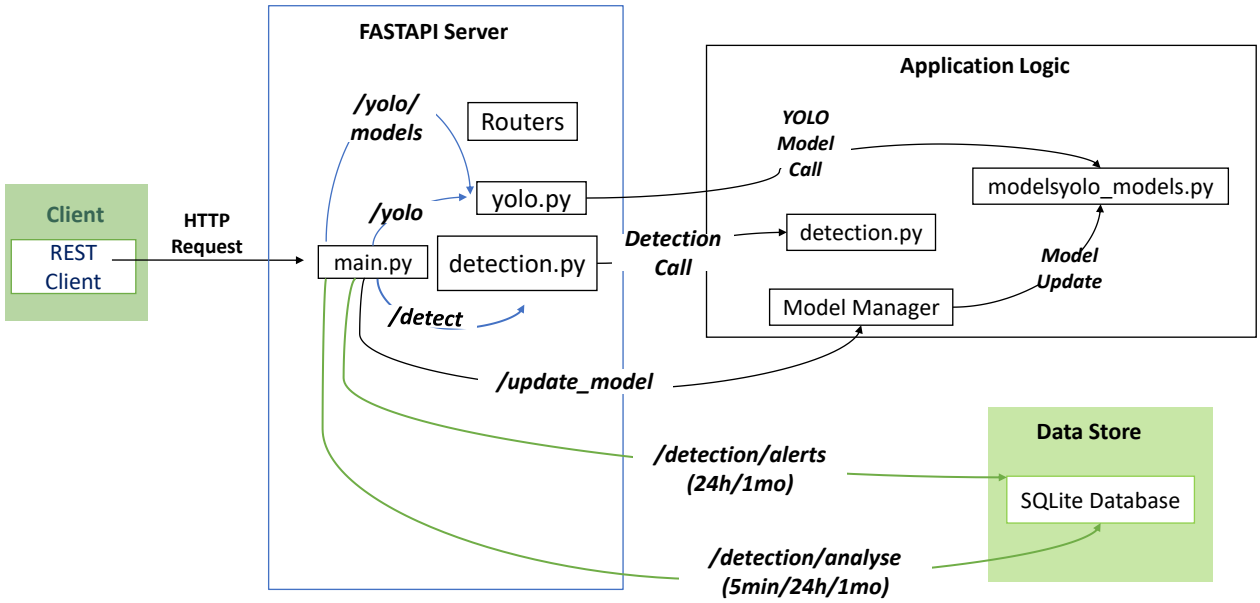


Figure 5: FastAPI Microservices Architecture for Real-time Hazard Detection

1. The Client represents the entry point into the system. It could be a web interface, mobile application, or REST client tools such as Postman or curl. These clients interact with the backend server by sending HTTP requests for tasks like submitting images for detection, updating models, or retrieving analytics.

2. The Server Layer operates as the command center of the system, built using FastAPI. The core routing begins with `main.py`, which directs requests to relevant handlers. Three primary routers include `yolo.py` for managing YOLO models and their metadata, `detection.py` for processing detection requests, and `/update_model` for managing updates to model weights. These routers expose endpoints such as `/yolo`, `/yolo/models`, `/detect`, `/detection/alerts`, and `/detection/analyze`, supporting metadata access, detection submissions, and real-time analytics.

3. The Application Logic encompasses the detection engines and control modules.

   `models/yolo_models.py` defines the YOLO model architecture, while `detector.py` handles the logic for processing detections.

4. Data Store is implemented using SQLite, archives detection logs, alerts, and analytics outputs. It supports the backend endpoints `/detection/alerts` and `/detection/analyze`, which allow users to extract insights over predefined periods.

5. The Alerting Mechanism is responsible for communicating safety risks in real time. When the system identifies a hazard, it instantly generates alerts both visually and digitally to notify workers or supervisors (see Fig. 5). This rapid alert delivery is vital for minimizing incident response time on active construction sites.

Table 1: API Component Overview and Interaction Flow

Component	**Role**	**Interactions**
REST Client	Sends HTTP requests	Communicates with FastAPI Server
main.py	Entry point, routes requests	Calls routers and application logic
yolo.py	Handles YOLO-related endpoints	Interacts with YOLO model logic
detection.py	Handles detection endpoints	Calls detector logic
detector.py	Performs detection using YOLO models	Uses models/yolo_models.py
models/yolo_models.py	Implements YOLO models	Used by detector and model manager
Model Manager	Updates YOLO models	Interacts with models/yolo_models.py
SQLite Database	Stores alerts and analysis results	Accessed via detection endpoints

4 Conclusion

This study introduces a modular and scalable real-time hazard detection framework tailored to the unique safety demands of construction environments. By integrating the YOLOv8n model with a FastAPI-based backend, the system provides fast, reliable, and high-fidelity detection suitable for deployment on-site. Challenges such as inconsistent lighting, environmental variability, and real-time alerting were systematically addressed through panoramic 360-degree data acquisition, edge-optimized model design, and a well-structured API interface. The pipeline, validated on annotated datasets from both standard and helmet-mounted 360-degree cameras, achieved precision and recall scores above 0.75. While performance was promising, thin or occluded hazards such as open wires remain particularly challenging; addressing false negatives in such safety-critical scenarios will therefore be a key focus of future work. Overall, the framework demonstrates strong potential as a critical layer of risk mitigation in digitally enabled construction sites. Future developments may extend alerting to mobile devices for supervisors, enabling on-site decision support. In addition, the modular API-based structure supports integration with dashboards, edge systems, and Industry Foundation Classes (IFC), linking hazard detection to BIM entities for automated and long-term safety management

References

[1] A. Pinto, I. L. Nunes, and R. A. Ribeiro, “Occupational risk assessment in construction industry – overview and reflection”, *Safety Science*, vol. 49, no. 5, pp. 616–624, 2011. DOI: 10.1016/j.ssci.2011.01.003

[2] Z. Zhou, Y. M. Goh, and Q. Li, “Overview and analysis of safety management studies in the construction industry”, *Safety Science*, vol. 72, pp. 337–350, 2015. DOI: 10.1016/j.ssci.2014.10.006

[3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection”, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[4] G. Jocher, *Yolov8: Next generation real-time object detector*, https://github.com/ultralytics/ultralytics, 2023.

[5] M. Zhang, T. Cao, and X. Zhao, “Applying sensor-based technology to improve construction safety management”, *Sensors*, vol. 17, no. 8, 2017. DOI: 10.3390/s17081841

[6] M. M. Alateeq, F. R. P.P., and M. A. S. Ali, “Construction site hazards identification using deep learning and computer vision”, *Sustainability*, vol. 15, no. 3, p. 2358, 2023. DOI: 10.3390/su15032358

[7] M. Arabshahi et al., “Review on sensing technology adoption in the construction industry”, *Sensors*, vol. 21, no. 24, 2021. DOI: 10.3390/s21248307

[8] C. Boje, A. Guerriero, S. Kubicki, and Y. Rezgui, “Towards a semantic construction digital twin: Directions for future research”, *Automation in Construction*, vol. 114, p. 103 179, 2020. DOI: 10.1016/j.autcon.2020.103179

[9] J. Karlapudi, K. Röhrl, M. Di’Pasquale, and K. Menzel, “Improving safety on construction sites by exploiting iot, digital twins, and data fusion”, in *Smart and Sustainable Collaborative Networks*, L. M. Camarinha-Matos, editor, Valencia: Springer, 2023. DOI: 10.1007/978-3-031-42622-3_46

[10] M. Sohan, S. Ram, and C. V. Reddy, “A review on yolov8 and its advancements”, in *Lecture Notes in Electrical Engineering*, 2024, pp. 529–545. DOI: 10.1007/978-981-99-7962-2_39

[11] M. Talib, A. H. Y. Al-Noori, and J. Suad, “Yolov8-cab: Improved yolov8 for real-time object detection”, *Karbala International Journal of Modern Science*, vol. 10, no. 1, 2024. DOI: 10.33640/2405-609X.3339

[12] D. Parkash, T. Kundu, and P. Kaur, “The rfid technology and its applications: A review”, *International Journal of Electronics and Computer Science Engineering*, vol. 1, no. 3, pp. 1447–1455, 2012.

[13] S. Hossen, “Analysis and mitigation of hazard situations and locations at construction sites”, Masterarbeit, TU Dresden, 2024.

A Digital Twin Prototype for a Model-Scale Robotic Vehicle: Architecture, Implementation, and Integration Perspectives

Fritz Trede[1], Mohammad Alallo[1], Youssef Aknin[1] and Vikram Hawladar[1]
[1]Institute of Construction Informatics, Applied Mathematics and Building Physics, Bauhaus-University Weimar, Weimar, Germany
E-mail(s): fritz.trede@uni-weimar.de, mohammad.alallo@uni-weimar.de, youssef.aknin@uni-weimar.de, vikram.hawladar@uni-weimar.de

Abstract: Digital twins (DT) offer promising tools for real-time monitoring, control, and simulation in civil infrastructure. While substantial progress has been made in DTs for stationary structures like bridges, less attention has been paid to mobile systems and their potential within DT ecosystems. This paper presents the system design of a DT prototype for a robotic vehicle, developed as part of a broader initiative exploring connected DT environments combining mobile and stationary systems. The system includes a modular object-oriented architecture, a finite-state-machine (FSM) controlled vehicle, and a user interface for real-time experiment setup and execution. We present the architectural design, identify limitations in implementation, and reflect on how the system supports integration between mobile and stationary DTs.

Keywords: Mobile System, Digital Twin, Finite State Machine, Object-Oriented Modelling, System of Systems

DOI: 10.18154/RWTH-CONV-254925. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

DTs are increasingly employed in civil infrastructure to enable real-time monitoring, predictive maintenance, and operational coordination. Emerging approaches emphasize connected DT ecosystems, where multiple independently managed DTs interact and share data across system boundaries. In such environments, mobile systems, such as vehicles, can act as dynamic data collectors, transmitting context information to stationary infrastructure DTs. Designing a mobile DT suitable for these roles remains an open challenge, particularly with respect to its communication logic, control architecture, and interoperability. In this paper, we present a prototype implementation of a digital twin for a model scale robotic vehicle, developed with the purpose of operating as a mobile asset within a DT ecosystem. The prototype was created for laboratory experiments and is designed as a platform to investigate synchronization, control, and data communication between a robotic vehicle DT and a stationary infrastructure DT - specifically a bridge.

2 Background and Related Work

A **DT** can be defined as a real-time, data-driven digital model of a physical asset, system, or process, continuously updated via sensors or data streams to reflect its current state and enable analysis, prediction, and optimization [1]. In civil engineering, DT technology has been implemented to support multiple use cases, such as structural health monitoring (SHM) or predictive maintenance, of physical assets [2], [3], [4]. For example, Costin et al. (2023) investigated the use of sensor readings (e.g., displacement, strain, vibration) in a bridge DT to detect damage or assess performance over time [3]. More recently, researchers have explored linking multiple DTs into interconnected systems. In this context, the term federated digital twin refers to loosely coupled DTs that share information across different systems or organizations to enable data validation and integrated analysis. Ramu et al. (2022) discuss how federated DTs in smart cities could allow independent systems (e.g. vehicles, roads, bridges) to collaborate by exchanging relevant data, without a single centralized model [5].

Applying the digital twin concept to **mobile systems** is a growing area of interest. To date, vehicle DT research has mostly centered on optimization of automotive performance and the predictive maintenance of the vehicles themselves (e.g. battery health in electric vehicles)[6]. In a civil engineering context, mobile systems can serve as moving probes in the built environment, collecting and transmitting data that complements stationary sensor networks. For instance, instrumented vehicles have been proposed for drive-by monitoring of bridges – measuring vibrations or deflections as they cross to infer structural conditions [7]. When such a vehicle's DT is synchronized with a bridge's DT, it enables cross-comparison of measurements, anomaly detection or validation of sensor readings. Despite this potential, the design of mobile DTs in federated settings is currently underexplored.

For the development of a DT, Jeong et al. propose a structured approach, consisting of sequential implementation stages. The first stage focuses on creating a digital representation of the target object. The second stage ensures that the states of the physical object and its digital counterpart are synchronized. In the third stage, simulations are used to analyze and predict real-world behavior based on virtual changes. The fourth and fifth stages emphasize the integration and interoperability of multiple DTs and the use of intelligent and autonomous technologies to manage their life cycles effectively [8]. Although integration and interoperability are formally addressed in later stages of DT development, data schemas and communication protocols must be considered early to ensure downstream compatibility. Further aspects to consider are data quality and transfer. For instance, post-processing might be required to ensure data is complete and free of noise. Finally, real-time exchange is further constrained by network limitations such as latency and low bandwidth [1], [9].

In summary, while the development of interconnected DT ecosystems is gaining traction, the integration between stationary and mobile DTs remains limited. This paper addresses this gap by designing and evaluating a DT prototype for a robotic vehicle. Following the first two implementation stages as described by Jeong et al., the prototype models a robotic vehicle as a digital replica over a sequence of discrete states, to enable real-time state synchronization between the physical and virtual entities. Although integration with a stationary system (i.e. a bridge) into a broader DT ecosystems is beyond the scope of this paper, the design choices will be reflected considering their extensibility and potential integration within a DT ecosystem. To this end, the research is guided by the following questions (RQ):

- **RQ1**: What architectural requirements must be addressed to develop a functional DT prototype for a robotic vehicle?
- **RQ2**: What implications do the design choices have for the future integration into a DT ecosystem?

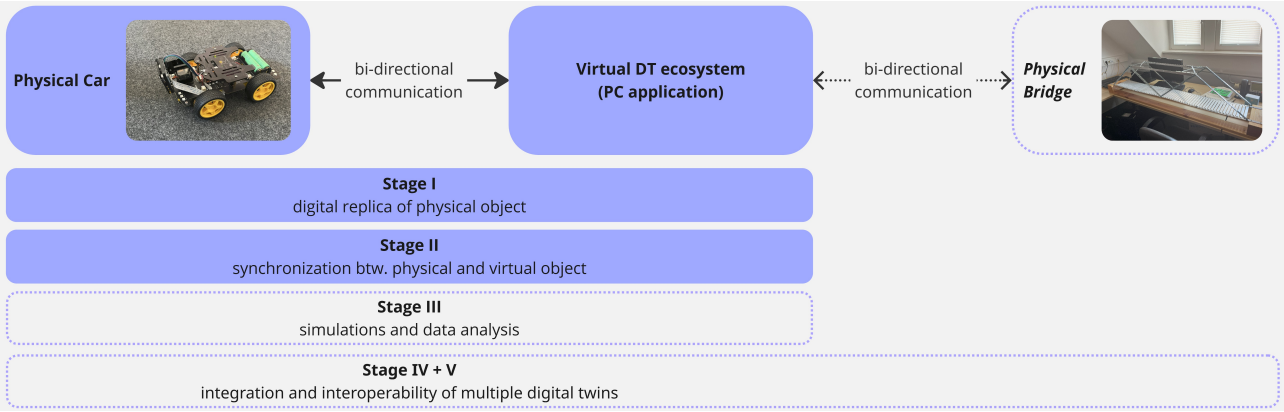


Figure 1: System Overview (Implementation Scope is marked purple)

3 System Overview

Beyond the intended integration into a broader DT ecosystem, the prototype was designed to support laboratory experimentation. To this end, key functional requirements entailed: the ability to define operational parameters, initiate experiment runs, and store operational parameters as well as aggregated data for subsequent analysis and future reference. To accommodate these requirements, the implementation focused on three core functional objectives:

1. **Real-Time Synchronization**: Build a virtual representation of the model car that reflects the car's physical state and behavior in real time.
2. **Data Logging**: Enable persistent storage of experimental setups and operational data for analysis and to prepare for integration with external systems.
3. **User-Driven Experimentation**: Provide an interactive graphical user interface (GUI) for researchers to configure experiment parameters, to start/stop runs, and to visualize operational data during the experiment.

To support these objectives, the prototype was developed using a modular architecture, seperating concerns related to vehicle control and synchronization (i.e. car communication interface), data management (i.e. database communication interface) and user interaction (i.e. GUI).

3.1 Information Flow and Message Schemas

To achieve the prototype's core objectives, the system relies on a clearly defined information flow between its components. This information flow is defined by two message schemas: experiment setup parameters and operational data. Both are serialized as key-value pairs using JavaScript Object Notation (JSON).

Table 1: SetupParameters and OperationalData JSON message schemas

SetupParameters *JSON*

Key	Value	Unit
Setup Name	String	
Target Speed	Integer	PWM
Running Time	Integer	Seconds
Update Interval	Integer	Seconds

OperationalData *JSON*

Key	Value	Unit
Motor Speed 1–4	Integer	PWM
Battery Capacity	Double	mAh
Battery Voltage	Double	V

Table 1 summarizes the specific parameters and data involved in both the setup and operation phases of the prototype. Figure 2 illustrates the information flow throughout the system. The setup parameters are specified via the GUI and transmitted to both the robotic car and a persistent database for future reference. The operational data is collected by the robotic vehicle during an experiment run and sent back through the car communication interface for visualization and storage.

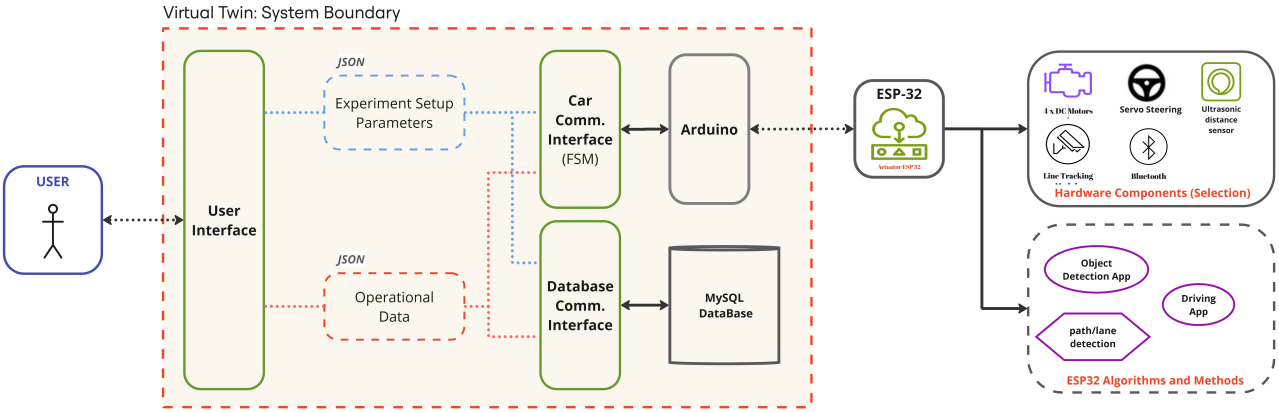


Figure 2: Information Flow

3.2 Hardware and Base Platform

The hardware platform for the system is a four-wheel drive robotic vehicle platform (i.e. Freenove robotic car kit) equipped with an ESP32 microcontroller as its main on-board computer. The ESP32 is a dual-core IoT-focused microcontroller with built-in WiFi and Bluetooth capabilities. All sensors and actuators on the car interface with the ESP32 via its general-purpose input/output pins. The car's hardware components include four direct current motors, a steering servo motor, and a suite of basic sensors. The DC motor speeds are regulated via pulse-width modulation (PWM), through varying width of voltage pulses. For this prototype, the sensor suite comprised an ultrasonic distance sensor (for obstacle detection), a line-tracking module (downward-facing sensors to follow a line on the floor, enabling an autonomous path-following mode), an infrared receiver, and built-in battery monitors. Dual 3.7V lithium batteries on the car power the motors and electronics [10].

3.3 Car Communication Interface

The Car Communication Interface handles the communication between the vehicles ESP32 microcontroller and the virtual twin via the Bluetooth Serial protocol. To manage and synchronize the car's behavior, we implemented an on-board FSM architecture. The ESP32's firmware, which was

programmed in C++ using the Arduino framework, is structured as an FSM that cycles through a set of four operational states. Each state is implemented with an associated handler function on the ESP32 (e.g. HandleInitialWaiting(), HandleRunning()) that encapsulates the logic for that phase. The FSM approach allows synchronizing the car's state and operations with the PC-side communication interface.

- `InitialWaiting`: The car is idle, awaiting a connection from the host PC. Upon powering up, the car stays in this state until a control link is established via Bluetooth. The ConnectToBluetooth() handler function sends a connection request to the ESP32, which sends a confirmation message and transitions to CommandWaiting.
- `CommandWaiting`: In this state the car is ready to receive experiment parameters, which are specified by the user via the GUI. The SendParameters() method transmits experiment parameters via JSON to ESP32. Upon receival, the ESP32 deserializes, sends confirmation, and transitions to WaitForRun.
- `WaitForRun`: In this state, the car has received a valid experiment setup and is ready to start the run. The SendRunMsg() method sends the run command, upon which the car will start transition to Running.
- `Running`: The car executes the experiment, carrying out the movement and recording pattern as specified. During this state, it continuously sends the operational data to the car communication interface via Bluetooth. The car remains in the Running state until the experiment duration elapses or a stop command is received, after which it transitions back to the `CommandWaiting` state for the next experiment.

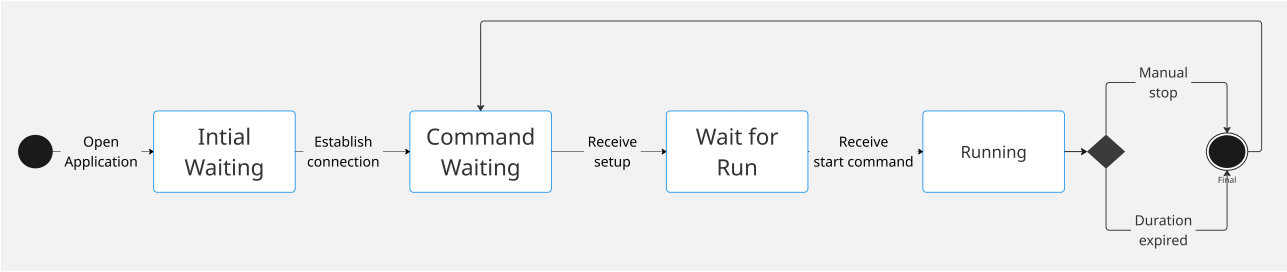


Figure 3: FSM States and Transition requirements

The program uses an event-driven approach to handle the receival of operational data from the car. Instead of continuously polling the system listens for events that are triggered when the car sends its operational data.

3.4 Database Communication Interface

The project uses a relational database system, MySQL, to store the aggregated data for experiment setups and operational data. The local database contains two tables `ExperimentSetup` and `OperationalData`, matching the contents of the JSON messages. The Database access is managed by a custom C# library class, which establishes a connection, and provides methods for inserting and fetching data stored in the database.

3.5 Graphical User Interface (GUI)

The GUI provides a structured workflow for users to configure experiment parameters, to start and stop experiment runs, and to visualize operational data during the experiment. The GUI was built following a Model-View-ViewModel design pattern, which separates the business logic of the FSM from the application interface. Over a sequence of Views, Users can create or select an experiment setup, according to the parameters expected by the car's FSM. The GUI includes validation to ensure that the required fields are provided, before serializing the user input into a JSON Object. After the user initiates the run, the interface displays the car's status and operational data in real time.

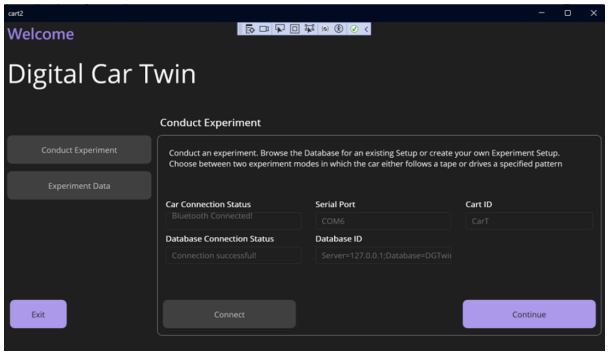


Figure 4: Main View

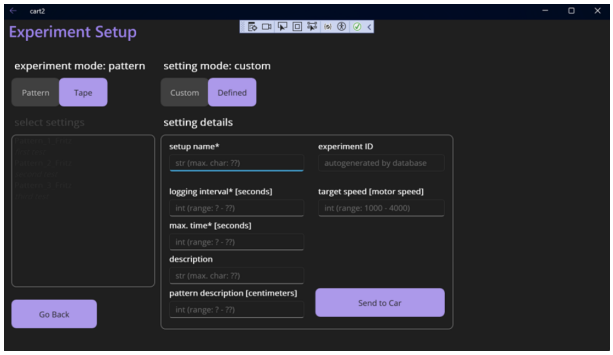


Figure 5: Experiment Setup View

4 Evaluation and Lessons Learned

The current implementation must be regarded as a prototype DT, since only stages 1 and 2 of Jeong et al.'s maturity model are realized — namely the creation of a digital representation and its synchronization with the physical vehicle [8]. Stages 3–5, which concern simulation-based prediction, interoperability with external DTs, and autonomous lifecycle management, remain unaddressed. Within this limited scope, communication is established via a Bluetooth serial interface and can be considered bidirectional: experiment setup parameters and run commands are sent from the GUI to the vehicle, while operational data such as motor speeds, battery voltage, and battery capacity are continuously transmitted back for visualization and storage. However, this bidirectional flow is restricted to parameter transmission and state reporting, without feedback mechanisms that would adapt system behavior. Furthermore, while operational data is persistently logged, no further analysis of these data streams has yet been implemented, leaving validation and performance assessment for future work.

Initial laboratory test runs confirmed that the prototype is functionally operable and capable of maintaining synchronization between the physical and virtual entities. However, systematic validation to assess the fidelity of the digital representation against the physical vehicle's behavior has not yet been conducted.

Having outlined the current state and limitations of the implementation, it is equally important to discuss practical findings and their implications for scalability and future integration.

The FSM architecture, which models the system through discrete states and simple transitions, has proven to be an effective approach for this resource-constrained mobile asset. Although its applicability

in more complex scenarios remains to be validated, the approach demonstrates strong potential in controlled laboratory settings, particularly where the primary role of the mobile system is the transmission of operational data. At present, the FSM design supports only unidirectional state transitions, extending it to include bidirectional transitions could provide greater flexibility in future development and integration. In terms of information flow, two message schemas were introduced, both using JSON as the serialization format. Within the current scope, the defined schemas are used to specify the interfaces within the system's modular architecture. However, in a system-of-systems context, additional mechanisms may be required to ensure clear data provenance and semantic alignment. Another critical aspect for such integration is the matchmaking of corresponding data points across system boundaries. In the current implementation, operational data is not timestamped, which limits temporal correlation between datasets. For future integration, timestamping will be essential to enable accurate data alignment. While PC-side timestamping is possible, it introduces uncertainty due to transmission latency. Therefore, device-side pre-transmission timestamping is preferred. This approach, however, requires synchronized device clocks to maintain temporal consistency across distributed components. With regard to scalability, it should be noted that while bandwidth demands for the experiment setup schema are expected to remain low, the operational data schema will likely require significantly more bandwidth as additional sensor data is integrated. This may exceed the capabilities of the current Bluetooth serial connection, which is limited in both bandwidth and range. As a result, future iterations may need to adopt alternative communication protocols with higher throughput. The use of a relational database for backend storage is generally appropriate; however, the current implementation using MySQL may become a bottleneck as system complexity increases. Issues such as reduced query performance, storage limitations, and restricted support for horizontal scaling could negatively impact future scalability. Finally, although data security and privacy are essential for real-world deployments, these issues fall outside the scope of this lab-scale study and remain subjects for future work.

5 Conclusion

This paper presented a prototype DT for a model-scale robotic vehicle, implementing stages 1 and 2 of Jeong et al.'s maturity model to achieve a synchronized digital representation and bidirectional communication via a Bluetooth interface. While initial test runs confirmed functional operation, systematic validation remains outstanding, and higher maturity stages such as simulation and integration have not yet been realized. The reflections on architecture, communication, and data management underline both the feasibility of the chosen design and the limitations that currently constrain scalability and interoperability. As such, the prototype serves as a foundation for future research towards more advanced and integrated DT ecosystems.

References

[1] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital twin: Enabling technologies, challenges and open research", *IEEE Access*, vol. 8, pp. 108 952–108 971, 2020. DOI: 10.1109/ACCESS.2020.2998358

[2] F. Jiang, L. Ma, T. Broyd, and K. Chen, “Digital twin and its implementations in the civil engineering sector”, *Automation in Construction*, vol. 130, p. 103 838, 2021. DOI: https://doi.org/10.1016/j.autcon.2021.103838 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580521002892

[3] A. Costin, A. Adibfar, and J. Bridge, “Digital twin framework for bridge structural health monitoring utilizing existing technologies: New paradigm for enhanced management, operation, and maintenance”, *Transportation Research Record*, vol. 2678, no. 6, pp. 1095–1106, 2023. DOI: 10.1177/03611981231208908

[4] C. Ye et al., “A digital twin of bridges for structural health monitoring”, Sep. 2019. DOI: 10.12783/shm2019/32287

[5] S. P. Ramu et al., “Federated learning enabled digital twins for smart cities: Concepts, recent advances, and future directions”, *Sustainable Cities and Society*, vol. 79, p. 103 663, 2022. DOI: https://doi.org/10.1016/j.scs.2021.103663 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210670721009264

[6] S. Verma, A. Sharma, B. Tran, and D. Alahakoon, “A systematic review of digital twins for electric vehicles”, *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 11, no. 5, pp. 815–834, 2024. DOI: https://doi.org/10.1016/j.jtte.2024.04.004 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S209575642400093X

[7] H. Shokravi, M. Vafaei, B. Samali, and N. Bakhary, “In-fleet structural health monitoring of roadway bridges using connected and autonomous vehicles’ data”, *Computer-Aided Civil and Infrastructure Engineering*, vol. 39, no. 14, pp. 2122–2139, 2024. DOI: https://doi.org/10.1111/mice.13180 eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/mice.13180. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/mice.13180

[8] D.-Y. Jeong et al., “Digital twin: Technology evolution stages and implementation layers with technology elements”, *IEEE Access*, vol. 10, pp. 52 609–52 620, 2022. DOI: 10.1109/ACCESS.2022.3174220

[9] W. Y. B. Lim et al., “Federated learning in mobile edge networks: A comprehensive survey”, *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 2031–2063, 2020. DOI: 10.1109/COMST.2020.2986024

[10] Freenove, *Freenove 4wd car kit for esp32 - tutorial*, https://github.com/Freenove/Freenove_4WD_Car_Kit_for_ESP32/blob/master/Tutorial.pdf, 2023.

Ontology-Augmented Multi-Agent Reinforcement Learning for Enhancing EV Charging Network Recovery

Mohamed Abdelfattah[1]
[1]Technische Universität Berlin, Berlin, Germany
E-mail(s): m.abdelfattah.ibrahim.mohamed@tu-berlin.de

Abstract: Electric Vehicle Charging Networks (EVCNs) face critical vulnerabilities during power outages, causing transportation disruption and grid instability. Our previous work introduced EVCAR, a Multi-Agent Reinforcement Learning (MARL) framework for post-outage recovery, but it operates with limited contextual awareness. This paper presents EVCAR-KG, extending our EVCAR system with a domain-specific knowledge graph that functions as a semantic digital twin. The framework enhances our original agents with ontology-augmented perception, knowledge-guided action formulation, and semantically augmented rewards. Experimental results show EVCAR-KG achieves 20 percent faster convergence (3,013 vs 3,764 episodes) and reduces recovery time from 43.0 to 40.0 hours compared to our baseline TD3-MADDPG implementation, with only 14 percent computational overhead. The complete implementation, including the ontology definition, agent integration code, and serialized OWL file, is publicly available at: https://doi.org/10.5281/zenodo.15824931

Keywords: Electric Vehicle Charging Networks Recovery, Prolonged Service Disruption, Grid Stability, Multi-Agent Reinforcement Learning, Knowledge Graphs

DOI: 10.18154/RWTH-CONV-254920. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction and Related Work

The rapid adoption of electric vehicles (EVs) presents significant challenges to the resilience of urban charging infrastructure. During power outages in densely populated districts, EV drivers often migrate en masse to neighboring areas, leading to congestion and voltage instability[1]. Events such as Berlin's 2019 blackout highlight the severity of such disruptions [2], increasing waiting times and overloading grid components.

The Electric Vehicle Charging Quality of Service Adaptive Recovery (EVCAR) system introduced a multi-agent deep reinforcement learning (MARL) approach to mitigate post-outage impacts [2]. EVCAR employs district-level agents adjusting charging rates and a central agent redistributing displaced EVs, achieving improved waiting times and grid stability compared to heuristic methods.

Building upon EVCAR, we investigate whether semantic knowledge integration can further enhance agent adaptability. Specifically, we explore ontology-driven knowledge graphs (KGs) to enrich agents'

spatial, infrastructural, and operational awareness. Recent research highlights ontology-enhanced reinforcement learning (RL) for better generalization, accelerated convergence, and reliable decision-making. One direction uses ontologies to directly guide agent behavior via embedded expert knowledge. For instance, the CORL framework integrated domain ontologies into classification tasks, turning RL into structured reasoning [3]. Similar frameworks appeared in cybersecurity with ontology-defined penetration strategies combined with RL and reasoning architectures [4], [5], allowing efficient action pruning and safety constraints.

Another approach preserves RL agent autonomy while enriching observations through ontology-derived contextual knowledge. Examples include the Automatic Goal Generation Model (AGGM), dynamically generating agent goals using ontology rules during state shifts [6], and the OnCertain model enhancing RL responses during uncertainty via ontology reasoning [7]. Furthermore, ontology reasoning effectively handled sensor data gaps in smart traffic systems by inferring missing states through upstream-downstream congestion relationships [8]. Such ontologies inject semantic awareness into agents, maintaining RL exploration while improving decision-making under partial observability.

Finally, ontology-driven knowledge graphs have enhanced multi-agent coordination. Manufacturing studies used shared ontologies for real-time adaptive scheduling, dynamically aligning task and resource understanding among agents [9], [10]. This concept of shared semantic understanding is particularly relevant for decentralized multi-agent systems like EVCAR, supporting agents' coordinated efforts toward maintaining grid stability and service equity.

This paper introduces EVCAR-KG, augmenting MARL with ontology-based knowledge to assess if structured knowledge enhances training and deployment adaptability. By aligning our ontology with schemas like EVKG [11], we ensure interoperability with real-world systems.

The primary research question is: *Can integrating a domain ontology as a knowledge graph improve MARL agents' learning efficiency and decision quality in EV charging network QoS recovery through embedded contextual knowledge?*

Our contributions include:

- Designing a domain-specific ontology capturing essential EV network relationships and operational states.
- Integrating the ontology with EVCAR MARL, enabling semantic agent observations.
- Evaluating improvements in learning efficiency and QoS recovery within a controlled EVCAR environment.

2 Methodology

2.1 System Architecture Overview

EVCAR-KG extends the baseline EVCAR system through three enhancements: (1) ontology-driven knowledge graph representing EV charging network structure/state, (2) semantic query interfaces for high-level contextual information, and (3) knowledge-informed decision mechanisms including action

masking and reward shaping. These extensions improve agent adaptability in variable scenarios, motivated by prior ontology-enhanced RL works [8], [5], [10]. Figure 1 shows the architecture.

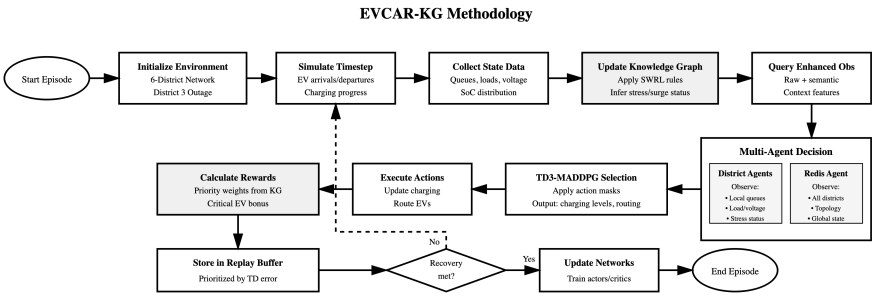


Figure 1: EVCAR-KG methodology flowchart showing the integration of ontology-driven knowledge graph with multi-agent reinforcement learning for EV charging network recovery

2.2 Ontology Architecture

The EVCAR-KG ontology provides a comprehensive semantic model of EV charging network recovery, structured in Web Ontology Language (OWL). It extends the EVKG schema [11] with domain-specific post-outage recovery concepts.

The ontology defines hierarchical classes capturing static infrastructure and dynamic operational states, distinguishing between infrastructure entities, operational events, and agent concepts (Figure 2).

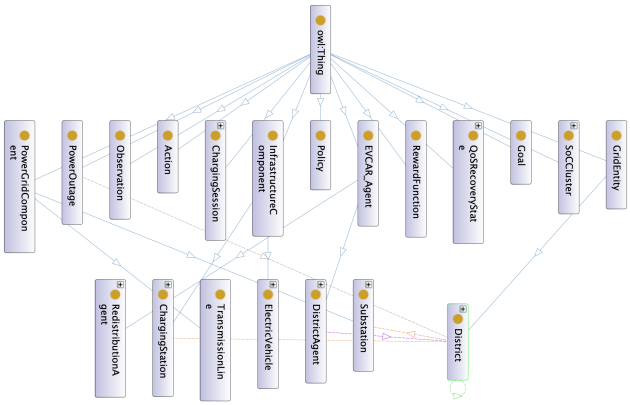


Figure 2: Ontology structure of EVCAR-KG in Protégé, showing key classes and semantic relationships relevant to EV charging network recovery.

Object properties establish network topology through spatial relationships (`locatedIn`, `hasNeighbor`), power relationships (`powers`, `suppliedBy`), and operational associations (`involvesVehicle`, `affectsDistrict`). Data properties capture both static characteristics (district types, battery capacity, station capacity) and dynamic state variables (queue lengths, voltage deviations, SoC levels, congestion states).

The ontology incorporates inference rules that derive high-level knowledge from raw data, implemented as Python functions. All rules are realised as lightweight Python callbacks; their outputs feed directly into masking and reward shaping (Table 1).

Table 1: Ontology-derived flags and their trigger conditions

Rule	Trigger condition (checked each step)	Flag/action
Grid stress	$\Delta V > 0.04$ p.u. & queue > 10 EV	`isUnderStress`=true
Surge	Neighbour outage & queue $\uparrow 50\%$ in < 1 h	`surgeStatus`=true
Priority vehicle	SoC $< 5\,\%$	`highPriority`=true (EV)
Invalid action	Route EV to stressed district	Mask / heavy penalty

The complete implementation, including the ontology definition, agent integration code, and serialized OWL file, is publicly available at: https://doi.org/10.5281/zenodo.15824931

2.3 Integration Methodology

2.3.1 State and action Augmentation

The knowledge graph fundamentally transforms how agents perceive their environment. Rather than processing raw numerical observations, agents receive semantically enriched state representations that incorporate both sensory data and inferred knowledge.

For district agents, the original observation vector from EVCAR, denoted as $o_i^d(t) \in \mathbb{R}^{15}$, undergoes augmentation with knowledge graph features to create an enhanced observation $\hat{o}_i^d(t) \in \mathbb{R}^{22}$. The augmentation process extracts seven additional features from the ontology:

$$\hat{o}_i^d(t) = [o_i^d(t), f_{type}^{res}, f_{type}^{com}, f_{type}^{ind}, f_{stress}, f_{surge}, f_{neighbors}^{out}, f_{cong}^{high}] \tag{1}$$

where f_{type}^{res}, f_{type}^{com}, and f_{type}^{ind} are binary indicators for district type (Residential, Commercial, Industrial), f_{stress} indicates whether the district is under stress according to inference rules, f_{surge} flags surge conditions, $f_{neighbors}^{out}$ counts neighboring districts with active outages, and f_{cong}^{high} indicates high congestion level.

For the redistribution agent, global state augmentation incorporates stress indicators and district types for all operational districts. The enhanced observation grows from 31 to 49 dimensions, adding three binary features per district (stress status, outage status, surge status). This comprehensive view enables system-wide coordination based on semantic understanding rather than purely numerical optimization.

The ontology also enables dynamic action space refinement through constraint propagation and safety rules. Unlike the original EVCAR where agents could select any action within their continuous or discrete space, EVCAR-KG implements ontology-based action masking.

For district agents, the action space modification occurs through a masking function $M : S \times A \to \{0, 1\}$ that filters available actions based on ontological constraints:

$$M(s, a) = \begin{cases} 0 & \text{if isUnderStress}(s) \wedge \text{increases_load}(a) \\ 1 & \text{otherwise} \end{cases} \tag{2}$$

This masking prevents agents from accepting additional EV load when their district is already stressed, except for critical SoC vehicles which maintain priority access. The implementation modifies the actor network output by applying the mask before action selection:

$$\pi_{masked}(a|s) = \frac{\pi(a|s) \cdot M(s,a)}{\sum_{a'} \pi(a'|s) \cdot M(s,a')} \tag{3}$$

For the redistribution agent, action masking operates at the district level. Districts flagged as stressed receive reduced allocation probabilities rather than complete exclusion, maintaining system flexibility while discouraging problematic routing:

$$p_{s,k}^{masked} = \begin{cases} 0.1 \cdot p_{s,k} & \text{if isUnderStress}(k) \\ p_{s,k} & \text{otherwise} \end{cases} \tag{4}$$

This soft masking approach, where stressed districts receive 10% of their original allocation probability, prevents complete isolation while strongly discouraging additional load.

2.3.2 Reward Function Enhancement

The integration of ontological knowledge enables sophisticated reward shaping that incorporates domain expertise and operational priorities. The enhanced reward function builds upon EVCAR's multi-objective structure while adding priority-based components.

For district agents, the original reward function from EVCAR undergoes modification to incorporate SoC cluster-specific weighting:

$$\hat{r}_i^d(t) = 0.8 \left(\sum_s w_s \cdot (SC_{i,s}(t) - QL_{i,s}(t) - WT_{i,s}(t)) - L_i(t) \right) + 0.2 \cdot r_{global}(t) \tag{5}$$

where $s \in \{critical, low, medium, high\}$ represents SoC clusters, and weights follow $w_{critical} = 2.0$, $w_{low} = 1.5$, $w_{medium} = 1.0$, $w_{high} = 0.8$. This prioritization ensures that agents learn to favor critical and low SoC vehicles, improving service quality for EVs most at risk of stranding.

The global reward component incorporates penalties for violating ontology-inferred constraints:

$$r_{global}(t) = r_{global}^{base}(t) - \alpha \sum_k \mathbb{1}[\text{isUnderStress}(k)] - \beta \sum_k \mathbb{1}[\text{surgeStatus}(k)] \tag{6}$$

where α and β are penalty coefficients for stressed and surge conditions respectively. This formulation encourages system-wide stress reduction beyond local optimization.

2.3.3 Learning Process Modifications

The knowledge graph integration fundamentally alters the learning dynamics of the TD3-MADDPG algorithm. Three key modifications enhance sample efficiency and convergence properties.

Prioritized Experience Replay incorporates ontological knowledge into transition prioritization. The priority calculation extends beyond TD-error to include semantic importance:

$$p_i = |\delta_i| \cdot \omega_{semantic} \tag{7}$$

where δ_i is the TD-error and $\omega_{semantic}$ is a weighting factor based on ontological properties:

$$\omega_{semantic} = \begin{cases} 2.0 & \text{if district stressed or during outage} \\ 1.5 & \text{if critical SoC ratio} > 0.1 \\ 1.0 & \text{otherwise} \end{cases} \tag{8}$$

This prioritization ensures that agents learn more frequently from critical situations, accelerating policy improvement in high-stakes scenarios.

Results

Training comparisons show that EVCAR-KG converged at episode 3,013, approximately 20% faster than the baseline TD3-MADDPG, which required 3,764 episodes (Figure 3). This improvement is attributed to the structured domain knowledge embedded in the ontology, which supports more efficient exploration through enriched state representations. However, this gain comes with an average computational overhead of around 14%, due to increased state dimensionality and ontology-related processing. The net result is a modest reduction in total training time.

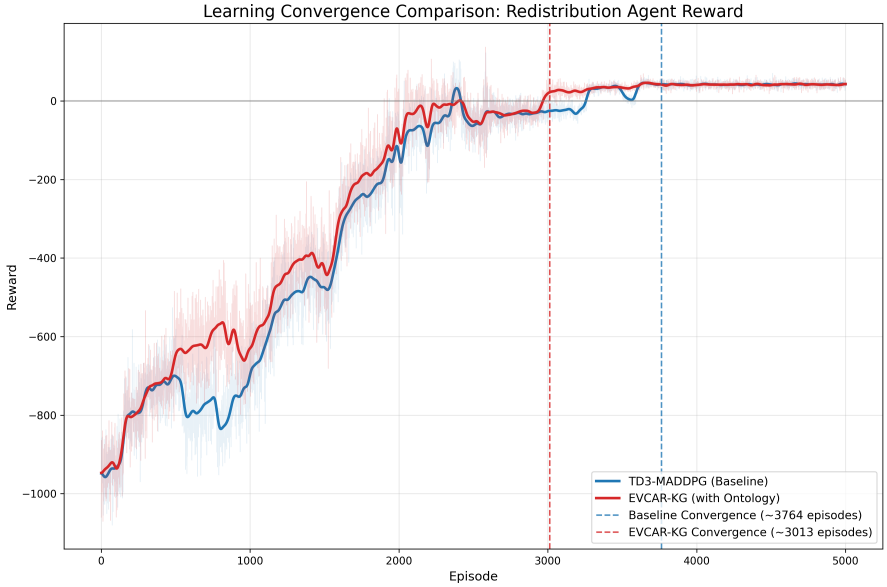


Figure 3: Comparison of agent convergence and policy evolution across episodes for base TD3-MADDPG and the integration of the ontology (EVCAR-KG).

Similarly, the district agents exhibited slightly faster convergence under EVCAR-KG, though the differences across most districts were not substantial. Districts 2 and 6, which were spatially removed from the outage, stabilized earlier, while the outage node (District 3) required more extended training, reflecting the higher complexity of local recovery under disrupted conditions.

Operationally, EVCAR-KG reduced post-outage queue lengths to acceptable levels within 40.0 hours, compared to 43.0 hours under TD3-MADDPG and significantly faster than traditional methods (60–102 hours). For vehicles in critical SoC states (<5%), both MARL approaches performed similarly, achieving safe levels within 11.9 and 12.0 hours respectively. System utilization recovered slightly faster under

EVCAR-KG, though again the margin was minimal. These results suggest that knowledge integration offers coordination advantages, particularly under spatially complex redistribution scenarios, while maintaining baseline performance across all recovery metrics.

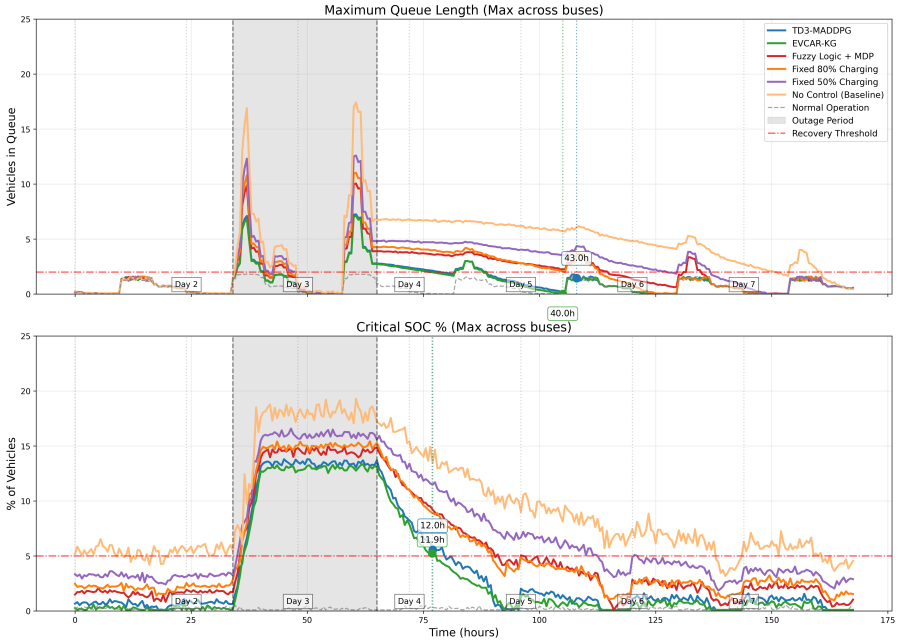


Figure 4: Comparison of user experience metrics (queue recovery and the percentage of critical SoC) across methods.

Conclusion

This paper presented EVCAR-KG, an ontology-enhanced extension of the EVCAR MARL framework for resilient recovery of electric vehicle charging networks. By integrating structured domain knowledge into the agent observation and action spaces, the system achieved slightly faster convergence and minor improvements in recovery metrics, particularly in queue length and system utilization. The performance for critical SoC management remained comparable to the baseline MARL implementation, indicating that the knowledge graph preserves coordination quality without introducing instability.

The observed gains came at the cost of approximately 14% increased computational effort. While this overhead is modest, it must be considered when deploying the system in resource-constrained environments. Nonetheless, the faster and more stable learning—particularly in complex or high-traffic districts—demonstrates the potential value of semantic knowledge integration in multi-agent recovery settings.

Future work will focus on optimizing the ontology to reduce runtime overhead, exploring the use of compressed or modular knowledge representations, and extending the system to larger urban environments with greater spatial and behavioral diversity. Additionally, real-world testing and hybrid architectures combining learning-based agents with rule-based decision modules may help strike a better balance between performance, interpretability, and scalability.

References

[1] D. Tang, X. Liu, T. Zhao, and Q. Wang, “Spatial–temporal qos assessment of the ev charging network considering power outages”, *Frontiers in Energy Research*, vol. 11, 2023. DOI: 10.3389/fenrg.2023.1112169

[2] M. Abdelfattah, N. Li, and T. Hartmann, “Evcar: Multi-agent reinforcement learning for electric vehicle charging network recovery”, in *EG-ICE 2025: International Workshop on Intelligent Computing in Engineering*, University of Strathclyde Publishing, 2025. DOI: 10.17868/strath.00093305

[3] B. Liu, L. Yao, Z. Ding, J. Xu, and J. Wu, “Combining ontology and reinforcement learning for zero-shot classification”, *Knowledge-Based Systems*, vol. 144, pp. 42–50, 2018. DOI: 10.1016/j.knosys.2017.12.022

[4] K. Qian, D. Zhang, P. Zhang, Z. Zhou, X. Chen, and S. Duan, “Ontology and reinforcement learning based intelligent agent automatic penetration test”, in *Proceedings of the 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, 2021, pp. 556–561. DOI: 10.1109/ICAICA52286.2021.9497911

[5] S. Sun, Y. Lu, D. Wu, and G. Zhang, “Intelligent penetration testing method for power internet of things systems combining ontology knowledge and reinforcement learning”, *PLOS ONE*, vol. 20, no. 5, e0323357, 2025. DOI: 10.1371/journal.pone.0323357

[6] S. Ghanadbashi and F. Golpayegani, “Using ontology to guide reinforcement learning agents in unseen situations”, *Applied Intelligence*, vol. 52, no. 2, pp. 1808–1824, 2022. DOI: 10.1007/s10489-021-02449-5

[7] S. Ghanadbashi, Z. Safavifar, F. Taebi, and F. Golpayegani, “Handling uncertainty in self-adaptive systems: An ontology-based reinforcement learning model”, *Journal of Reliable Intelligent Environments*, vol. 10, no. 1, pp. 19–44, 2024. DOI: 10.1007/s40860-022-00198-x

[8] J. Guo, S. Ghanadbashi, S. Wang, and F. Golpayegani, “Urban traffic signal control at the edge: An ontology-enhanced deep reinforcement learning approach”, in *Proceedings of the 2023 IEEE International Conference on Intelligent Transportation Systems (ITSC)*, 2023, pp. 6027–6033. DOI: 10.1109/ITSC57777.2023.10422573

[9] S. Qu, J. Wang, S. Govil, and J. O. Leckie, “Optimized adaptive scheduling of a manufacturing process system with multi-skill workforce and multiple machine types: An ontology-based, multi-agent reinforcement learning approach”, *Procedia CIRP*, vol. 57, pp. 55–60, 2016. DOI: 10.1016/j.procir.2016.11.011

[10] Z. Qin and Y. Lu, “Knowledge graph-enhanced multi-agent reinforcement learning for adaptive scheduling in smart manufacturing”, *Journal of Intelligent Manufacturing*, 2024. DOI: 10.1007/s10845-024-02494-0

[11] Y. Qi, G. Mai, R. Zhu, and M. Zhang, “Evkg: An interlinked and interoperable electric vehicle knowledge graph for smart transportation system”, *Transactions in GIS*, vol. 27, no. 4, pp. 949–974, 2023. DOI: 10.1111/tgis.13064

Knowledge-Enhanced Cobot Reasoning Assistant for Robot Manipulation

Ruibo Tang[1], Clara-Larissa Lorenz[1], Jérôme Frisch[1] and Christoph van Treeck[1]
[1]Institute of Energy Efficiency and Sustainable Building (E3D), RWTH Aachen University, Aachen, Germany
E-mail(s): ruibo.tang@rwth-aachen.de, lorenz@e3d.rwth-aachen.de, frisch@e3d.rwth-aachen.de, treeck@e3d.rwth-aachen.de

Abstract: The construction industry is exploring collaborative robots (cobots) and sensing technologies to support construction automation. With the emergence of large language models (LLMs), their integration with cobots could introduce a new paradigm in construction, where language serves as a direct interface for guiding and coordinating robotic assistance. However, language instruction control of cobots in dynamic environments remains challenging: (a) LLMs may generate actions that are plausible but not feasible, (b) long-horizon tasks may require interactive human feedback during execution, and (c) developing high-quality domain-specific knowledge bases remains costly and time-consuming. In this work, we present *Knowledge-Enhanced Cobot Reasoning Assistant* (***KECRA***), an LLM and Visual Foundation Models (VFMs) based framework. ***KECRA*** enables intuitive dialogue-based interaction with human operators, reducing the need for expert knowledge in programming or task specification. VFMs act as the eyes of ***KECRA***, perceiving target objects and capturing real-time spatial context to inform planning and execution. Core reasoning uses a state-graph agent (*Instruction* → *Plan Sub-tasks* → *Confirm Sub-tasks* → *Generate Path* → *Confirm Path*) with rewindable human checkpoints at key nodes for plan correction. All interactions are automatically accumulated into a knowledge graph, enabling LLMs retrieval-augmented reasoning. We validate ***KECRA***'s practicality and effectiveness in tabletop manipulation experiments.

Keywords: Collaborative Robots, Large Language Model, Multi-Modal, AI Agent, Knowledge Graph

DOI: 10.18154/RWTH-CONV-254921. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Robotic applications in construction aim to address labor shortages and improve productivity, with collaborative robots (cobots) enabling safe human–robot interaction [1], [2]. To expand their flexibility and usability, more recent studies have focused on using LLMs, which offer zero-shot learning, commonsense reasoning, and contextual understanding capabilities, for robot manipulation [3], [4], [5], [6]. However, language instruction control of cobots in unstructured environments such as construction sites remains challenging: (1) LLMs may generate action sequences that are commonsensical but physically infeasible in the current scene due to hallucination [7], [8]; (2) end-to-end LLM planning

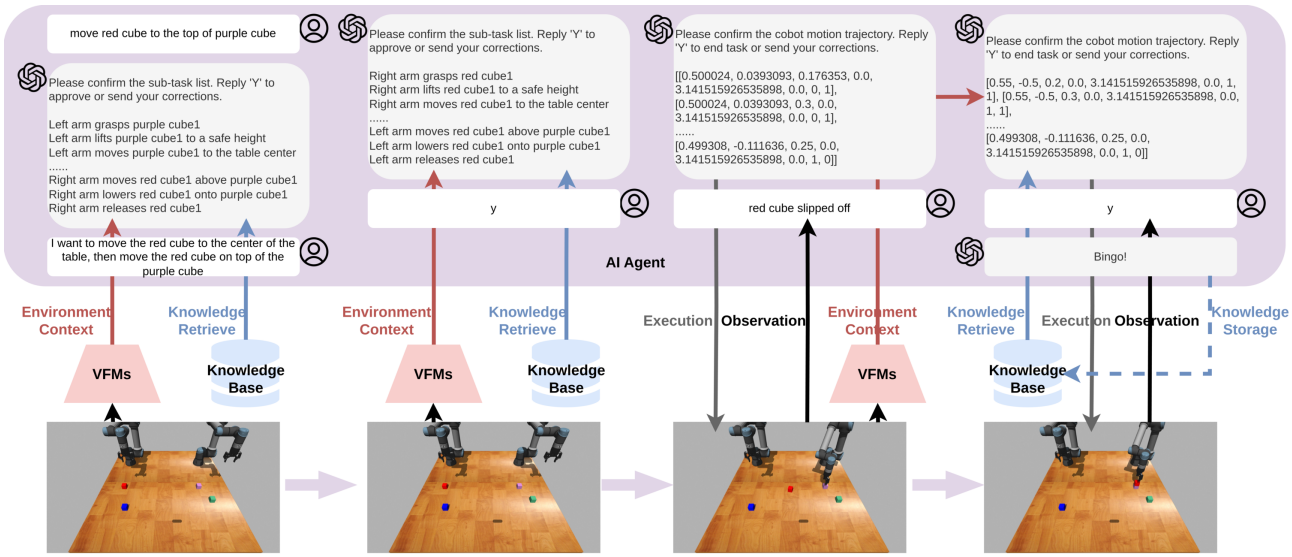


Figure 1: Overview of KECRA with example task "move red cube to the top of purple cube".

pipelines lack mid-execution feedback mechanisms, preventing timely error correction and introducing safety hazards [7], [9]; and (3) Retrieval-Augmented Generation (RAG) systems struggle to interpret implicit operational preferences in instructions [4]. A representative case is spatial command ambiguity, such as "place the object beside another", does "beside" mean to the left or to the right?

To address these challenges, we propose ***KECRA***, *Knowledge-Enhanced Cobot Reasoning Assistant*, which integrates a transparent loop state graph with human-in-the-loop (HIL) checkpoints to enable controllable and interpretable plan execution, and a structured knowledge graph that incrementally stores prior task interactions to support RAG. We show the code of ***KECRA*** at https://github.com/RWTH-E3D/KECRA. For example, when given the instruction “move the red cube to the top of the purple cube” (Figure 1), the high-level task planner integrates environmental context with prior task knowledge and decomposes the command into sub-tasks for operator approval. The operator revises the plan, prompting the agent to regenerate sub-tasks using the updated context and retrieved feedback examples. After validation, a low-level pose planner generates the trajectory and sends it to the robot arms for execution. When a slip occurs, the agent replans with the new context, completes the task, and logs the entire interaction to the knowledge base for future use.

The main contributions are: (1) a novel cobot assistant framework ***KECRA***, which integrates VFMs, a state-graph planner with human verification, and a self-enhancing knowledge graph to improve task success in dual-arm tabletop manipulation; (2) a pipeline that automatically converts human–AI interaction histories into structured knowledge graphs, enhancing planning efficiency while sustaining performance for complex instructions.

2 Related Work

LLM-powered robotic manipulation: LLMs with hundreds of billions of parameters, such as GPT-4o [10], have advanced language-conditioned robotic manipulation by enhancing in-context planning and reasoning capabilities [3]. Recent research on robotic manipulation leverages these foundation

models in three main directions: (1) *zero-shot action sequence generation*, where LLMs infer task steps directly from natural language [11]; (2) *HIL adaptation*, allowing real-time feedback to refine plans during execution [12], [13]; and (3) *direct code synthesis*, replacing predefined skill libraries with executable programs [4], [5]. In addition, a set of auxiliary modules are often integrated into LLM-based planners to ground task objectives and prevent hallucinations: (1) reward mechanisms to align generated plans with physical and safety constraints [7], [14]; (2) conformal prediction frameworks to resolve ambiguity through structured multiple-choice reasoning [13]; and (3) vision-language models (VLMs) to provide environmental grounding [5], [15]. However, most prior works implement either high-level task planning or low-level code generation in isolation [16], lacking ***KECRA***'s dual-phase integrated planning that ensures long-horizon tasks are executed in alignment with operator preferences. Moreover, corrective feedback is typically handled as ad hoc interventions [17], rather than through ***KECRA***'s rewindable state-graph workflow, which enables transparent backtracking to specific reasoning nodes when errors are detected. While VLMs and reward functions in previous works are often used as post hoc success detectors [18], ***KECRA*** leverages them as active causal context providers: real-time semantic perceptions (e.g., object dimensions, spatial relations) directly condition LLM reasoning, preventing physically infeasible plans and informing context-aware replanning.

Knowledge base for LLMs reasoning: LLMs with RAG rely on domain-relevant knowledge bases to retrieve indexed text corpora as contextual examples [19], typically using sentence embeddings for similarity matching [20]. In language-conditioned robotic manipulation, especially under multi-turn interaction scenarios, the design of the knowledge base fundamentally shapes how LLM outputs translate into robotic control [8], [21]. Most prior works ground LLM reasoning on one of the following knowledge sources: (1) static few-shot examples embedded in the prompt [5], [7], [21]; (2) rule summaries distilled from multi-turn feedback [6], [22]; (3) structured symbolic programs or task graphs that the LLMs can inspect and modify [4], [8]; and (4) skill-level or affordance databases indexed via embeddings for feasibility checking [7], [13]. However, these pipelines often lack causal traceability and spatiotemporal constraints, limiting their ability to resolve ambiguities like spatial relations [6], [23]. In contrast, ***KECRA*** introduces a structured knowledge graph that captures prior operator interactions, encoding instructions, sub-task sequences, environment context and correction paths. This design enables LLMs to retrieve and reason over causally coherent examples, resolving spatial ambiguities and supporting preference-aware adaptation with higher consistency.

3 Methods

We address two key challenges in language-conditioned cobot control: (1) how to align black-box LLMs with precise, step-by-step cobot execution to ensure transparency and controllability, and (2) how to make the system learn from history experience to human preferences for faster and more accurate task completion. To tackle this, we design **KECRA**, a unified framework that integrates three components: (1) VFMs for real-time semantic scene understanding; (2) a Graph agent for structured reasoning and HIL planning; and (3) an online knowledge graph that stores interaction history as causally consistent subgraphs to support retrieval-augmented reasoning. Together, they form a perception–planning–memory loop for transparent, adaptive cobot manipulation.

3.1 Capturing Environment Context with VFMs

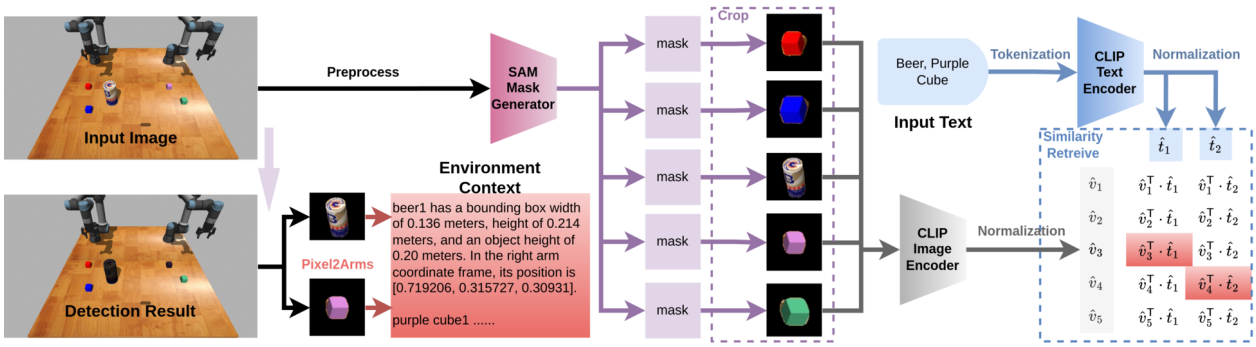


Figure 2: The environment context capturing with example input "Beer, Purple Cube".

During detection (Figure 2), object names are tokenized and encoded with the pretrained *Contrastive Language–Image Pretraining* (CLIP) [24] text encoder to obtain normalized text feature vectors $\hat{\mathbf{t}}_k \in \mathbb{R}^d$. In parallel, pretrained *Segment Anything Model* (SAM) [25] generates segmentation masks from RGB images, which are cropped and encoded by the CLIP image encoder to produce image features $\hat{\mathbf{v}}_i \in \mathbb{R}^d$. Recognition is performed by cosine similarity followed by softmax normalization:

$$Score(i) = \frac{\exp\left(\hat{\mathbf{v}}_i^\top \hat{\mathbf{t}}_k\right)}{\sum_{j=1}^{M} \exp\left(\hat{\mathbf{v}}_j^\top \hat{\mathbf{t}}_k\right)}, \tag{1}$$

where $Score(i)$ is the probability that candidate i belongs to class k. A match is accepted if $Score(i) > 0.2$, a threshold selected from validation to ensure reliable discrimination.

For each matched object, its pixel center (x_c, y_c) is derived from the bounding box, and height is estimated from the depth value at (x_c, y_c). Using camera intrinsics and depth, 2D coordinates are projected into 3D positions and dimensions, then transformed into each arm's coordinate frame by extrinsic calibration. Finally, this context is structured into a natural language environment description, as shown in Figure 2.

3.2 Planning with Graph Agent

KECRA's planning module employs a state graph agent built with LangGraph [26], enabling structured reasoning with human oversight. Unlike linear prompt chains, the agent uses explicit Chain-of-Thought (CoT) decomposition to break high-level instructions into verifiable steps, improving interpretability and reducing hallucinations. HIL checkpoints allow operators to inspect and revise plans at each stage for safe mid-execution correction.

The planner consists of seven nodes: *Start*, *Plan Tasks*, *Confirm Tasks*, *Generate Poses*, *Confirm Poses*, *End*, and the error-handling node *Pose Fix* (Figure 3). At *Plan Tasks*, the LLM integrates the instruction with environmental context and similar past episodes, producing a sub-task list ($tasks$). This list is verified at *Confirm Tasks*; if rejected, operator feedback ($feedback$) triggers iterative replanning.

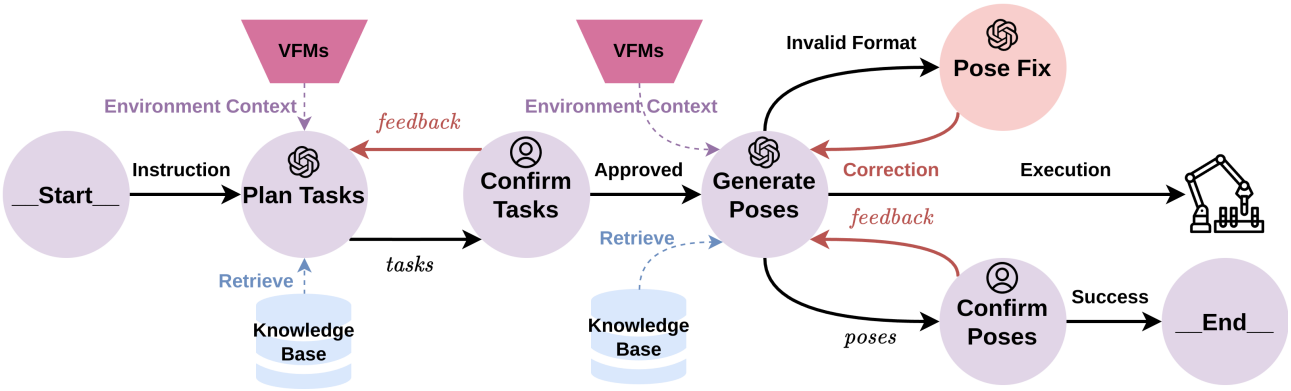


Figure 3: KECRA's state graph planning workflow.

Once approved, *Generate Poses* converts $tasks$ into pose sequences $[x, y, z, roll, pitch, yaw, arm, gripper]$, stored as $poses$. Invalid outputs are redirected to *Pose Fix* for correction. Accepted $poses$ are streamed to the robot arms and verified at *Confirm Poses*. If execution fails (e.g., object slippage), updated context and retrieved correction episodes guide replanning. The loop continues until poses are successfully executed, after which the graph transitions to *End*.

3.3 Memory Enhancement with Knowledge Graph

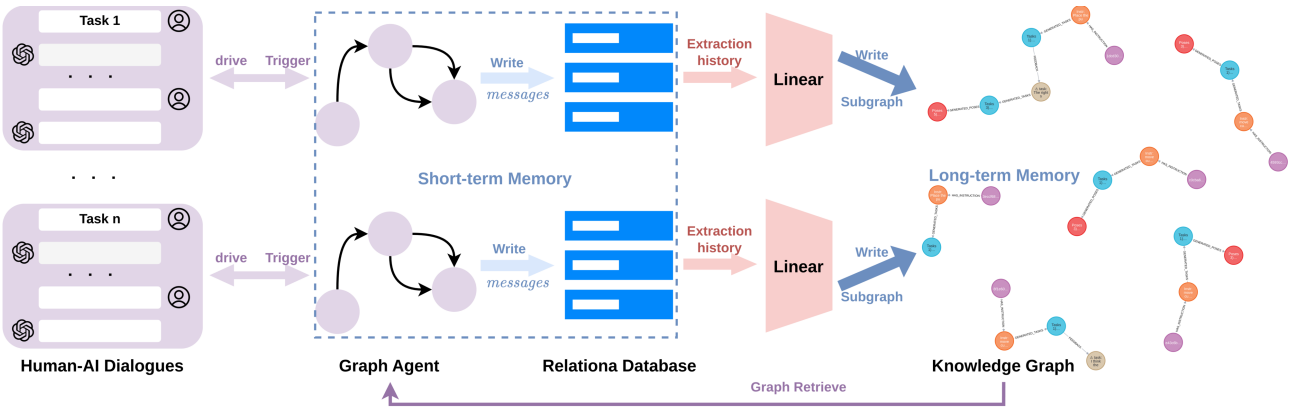


Figure 4: Overview of KECRA's memory architecture.

KECRA's memory system (Figure 4) has two components: (1) **short-term memory**, which stores ongoing interaction context (instructions, task lists, poses, feedback) in the state variable $messages$ and persists it via checkpointing; and (2) **long-term memory**, which transforms completed task histories into a knowledge graph in a graph database, linking instructions, sub-tasks, poses, and feedback.

When the state graph reaches *End*, each episode is unfolded into a causal subgraph (e.g., *Instruction* $\rightarrow$ *TaskList* $\rightarrow$ *PoseList* with feedback links) and added to the knowledge graph. This graph supports RAG through three mechanisms: (1) **Plan Tasks**, where the agent retrieves instructions similar to the operator's input and uses the final task and pose lists of mature episodes (with few corrections) as few-shot references; (2) **Confirm Tasks**, where task revisions trigger a search for past episodes with similar corrections, and the before/after task spans are supplied as guidance for replanning; and (3)

Confirm Poses, where execution failures prompt retrieval of prior pose-level corrections, using the transition from incorrect to corrected poses as regeneration examples.

4 Experiment

4.1 Experiment Setup

Simulation Environment: Experiments are conducted in Gazebo with two UR5e arms and four colored cubes. Two cubes are reachable only by the left arm and two only by the right, creating scenarios for both single-arm and coordinated dual-arm manipulation.

Language Models: KECRA uses GPT-4o for reasoning and *text-embedding-3-small* [27] for similarity retrieval in RAG, both via the OpenAI API.

Task Categories: We evaluate three categories: (1) single-arm placement, (2) cross-arm handover and placement, and (3) long-horizon instructions. Each includes 100 trials, with five linguistic variants per instruction to test robustness.

Baselines and Ablations: We compare KECRA to: (1) a prompt-only baseline using GPT-4o, where the prompt is optimized with demonstrations extracted from KECRA's own successful dialogues to improve fairness, but without feedback or memory mechanisms, and (2) an ablation without the knowledge graph, isolating its contribution.

Metrics: Performance is measured by **task success rate** and **average execution steps**. The prompt-only baseline completes in one step, while KECRA requires at least four (*Plan* $\rightarrow$ *Confirm Tasks* $\rightarrow$ *Generate Poses* $\rightarrow$ *Confirm Poses*); each additional human feedback adds two steps.

4.2 Experiment Results

From Table 1 we observe KECRA-full delivers the highest reliability, achieving 95%, 89%, and 77% success on the single-arm, cross-arm and long-horizon sets, respectively. The checkpointed state graph restricts the execution flow to at least four steps, while feedback loops raise the average to 4.38–6.74, but they prevent most infeasible plans.

Removing the knowledge graph lowers success by 9 percentage points on average and increases the step cost to 6.78–10.13. Without retrieval, the agent requires additional feedback rounds to eliminate hallucinated sub-tasks or kinematically invalid poses, confirming the benefit of experience reuse.

The prompt-only baseline succeeds in at most 71% of single-arm cases and almost never solves long-horizon tasks; its single step planning precludes intermediate correction. For example, in a long-horizon scenario ("stack the green cube on the blue, then move it to the left of the purple cube"), the baseline often failed to execute the second step because the initial plan ignored the change in object location after the first action, leading to incorrect grasp positions. KECRA succeeded by incorporating mid-execution corrections and retrieving prior episodes, enabling it to adjust the plan and trajectory accordingly. These results demonstrate that: (1) a transparent loop state graph with human checkpoints markedly improves task completion, and (2) the automatically accumulated knowledge graph accelerates convergence by supplying specific examples for retrieval-augmented reasoning.

Table 1: Success rate (%) and average execution steps.

	Success ↑			Steps ↓		
Task	Single	Cross	Long	Single	Cross	Long
KECRA (full)	**95**	**89**	**77**	**4.38**	**4.99**	**6.74**
KECRA – w/o KG	86	78	55	6.78	7.19	10.13
Prompt-only baseline	71	28	5	1.00	1.00	1.00

5 Conclusion

We develop and validate KECRA as a transparent controller for cobot operations. It transforms black-box LLMs into transparent cobot controllers through synergistic integration of visual perception (VFMs), a rewindable state-graph agent for human verified planning, and an automatically constructed knowledge graph capturing causal interaction histories. Validation in dual-arm tabletop tasks demonstrates that KECRA significantly outperforms prompt-only baseline. In complex cross-arm and long-horizon scenarios, KECRA maintains high reliability while substantially reducing correction cycles.

References

[1] T. Wang, C. Mao, B. Sun, and Z. Li, “Genealogy of construction robotics”, *Autom. Constr.*, vol. 166, p. 105 607, 2024.

[2] M. Schöberl, A. Huber, S. Kreppold, J. Dirnaichner, S. Kessler, and J. Fottner, “Cobot uptake in construction: Embedding collaborative robots in digital construction processes”, *Constr. Robot.*, vol. 7, no. 1, pp. 89–103, 2023.

[3] H. Zhou et al., *Bridging language and action: A survey of language-conditioned robot manipulation*, arXiv preprint arXiv:2312.10807, 2025.

[4] J. Liang et al., “Code as policies: Language model programs for embodied control”, in *Proc. of the ICRA*, 2023, pp. 9493–9500.

[5] S. Huang, Z. Jiang, H. Dong, Y. Qiao, P. Gao, and H. Li, *Instruct2act: Mapping multi-modality instructions to robotic actions with large language model*, arXiv preprint arXiv:2305.11176, 2023.

[6] L. Zha et al., “Distilling and retrieving generalizable knowledge for robot manipulation via language corrections”, in *Proc. of the ICRA*, 2024, pp. 15 172–15 179.

[7] M. Ahn et al., *Do as i can, not as i say: Grounding language in robotic affordances*, arXiv preprint arXiv:2204.01691, 2022.

[8] I. Singh et al., “Progprompt: Generating situated robot task plans using large language models”, in *Proc. of the ICRA*, 2023, pp. 11 523–11 530.

[9] M. Zhang, R. Xu, H. Wu, J. Pan, and X. Luo, “Human–robot collaboration for on-site construction”, *Autom. Constr.*, vol. 150, p. 104 812, 2023.

[10] OpenAI, *Gpt-4o system card*, arXiv preprint arXiv:2410.21276, 2024.

[11] W. Huang, P. Abbeel, D. Pathak, and I. Mordatch, “Language models as zero-shot planners: Extracting actionable knowledge for embodied agents”, in *Proc. of the ICML*, PMLR, 2022, pp. 9118–9147.

[12] W. Huang et al., “Inner monologue: Embodied reasoning through planning with language models”, in *Proc. of the CoRL*, 2022.

[13] A. Z. Ren et al., “Robots that ask for help: Uncertainty alignment for large language model planners”, in *Proc. of the CoRL*, 2023.

[14] W. Yu et al., “Language to rewards for robotic skill synthesis”, in *Proc. of the CoRL*, 2023.

[15] S. Mirchandani et al., “Large language models as general pattern machines”, in *Proc. of the CoRL*, 2023.

[16] C. Rivera et al., *Conceptagent: Llm-driven precondition grounding and tree search for robust task planning and execution*, arXiv preprint arXiv:2410.06108, 2024.

[17] S. Izadi and M. Forouzanfar, “Error correction and adaptation in conversational ai: A review of techniques and applications in chatbots”, *Ai*, vol. 5, no. 2, pp. 803–841, 2024.

[18] Z. Liu, A. Bahety, and S. Song, “REFLECT: Summarizing robot experiences for failure explanation and correction”, in *Proc. of the CoRL*, 2023.

[19] P. Lewis et al., “Retrieval-augmented generation for knowledge-intensive nlp tasks”, in *Proc. of the NeurIPS*, 2020, pp. 9459–9474.

[20] Z. Li, Z. Wang, W. Wang, K. Hung, H. Xie, and F. L. Wang, “Retrieval-augmented generation for educational application: A systematic survey”, *Comput. Educ.: Artif. Intell.*, vol. 8, p. 100 417, 2025.

[21] C. H. Song, J. Wu, C. Washington, B. M. Sadler, W.-L. Chao, and Y. Su, “Llm-planner: Few-shot grounded planning for embodied agents with large language models”, in *Proc. of the ICCV*, 2023, pp. 2998–3009.

[22] S.-Q. Yan, J.-C. Gu, Y. Zhu, and Z.-H. Ling, *Corrective retrieval augmented generation*, arXiv preprint arXiv:2401.15884, 2024.

[23] Y. Yuan et al., *From seeing to doing: Bridging reasoning and decision for robotic manipulation*, arXiv preprint arXiv:2505.08548, 2025.

[24] A. Radford et al., “Learning transferable visual models from natural language supervision”, in *Proc. of ICML*, PmLR, 2021, pp. 8748–8763.

[25] A. Kirillov et al., “Segment anything”, in *Proc. of the ICCV*, 2023, pp. 4015–4026.

[26] J. Wang and Z. Duan, *Agent ai with langgraph: A modular framework for enhancing machine translation using large language models*, arXiv preprint arXiv:2412.03801, 2024.

[27] OpenAI, *Openai embeddings api*, 2024. [Online]. Available: https://platform.openai.com/docs/guides/embeddings

A system for querying spatial relations between technical drawings and geometric data using SPARQL

Fabian Rolke[1]
[1]Chair of Design Computation (DC), RWTH Aachen University, Aachen, Germany
E-mail(s): rolke@dc.rwth-aachen.de

Abstract: The increasing size and heterogeneity of construction related project databases demands significant effort from a limited number of data experts in the architecture, engineering, construction, and operation (AECO) industry. Advancements in technologies for capturing building data necessitate ongoing development in transformation, integration, and interoperability. A key objective is linking heterogeneous building data through their respective spatial alignment. While several RDF and SPARQL extensions support two-dimensional relational querying, there is a notable lack of technologies capable of processing three-dimensional data. Technical drawings are still one of the most common mediums for transferring geometrical data between stakeholders due to their abstract, reduced representational quality, high information density, and extensive use on construction sites. However, they often exist in non-machine-readable formats that hinder queries on depicted data such as location, rotation, scale relative to built assets, and individual annotations. This paper presents an approach to link and query topological, directional, and metric relationships among 2D plans and other data types by storing their three-dimensional spatial information as RDF. A spatial calculus identifies correlations between selected entities via SPARQL queries. Results are written back in RDF to enrich existing databases or provide specific answers about entity relationships.

Keywords: RDF, SPARQL, technical drawing, spatial querying, geometry processing

DOI: 10.18154/RWTH-CONV-254922. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

One of the biggest tasks in the architecture, engineering, construction, and operation industry is addressing the challenges of closing the gap between the real and digital world. For this purpose, conceptualizing spatial relations between entities is just as significant as defining their geometrical representations. All spatial relations, whether real or digital, are classified as either topological, directional, or metric [1] [2]. Randell et al. provide an overview of the base topological relations: disjoint, meet, overlap, equal, coveredBy, inside, covers, and contains [3]. Building on this work, the Relative Location (RELOC) Ontology by Göbels and Beetz allows for the conversion of natural-language-based location descriptions into RDF-structured spatial relationships, thus making them

machine-readable. Topological relations such as meet, intersect, and containedIn are combined with directional relations such as front, right, or vertical center, allowing for spatial specification in RDF without the immediate need for vector data serializations [4]. Göbels and Schulz further outlined the need for a spatial vocabulary enabling classification of different space types when working with large, heterogeneous datasets [5]. An Entity Space, defined as a singular, spatial "thing", is a subclass of Asset Space which is any kind of overall reference model of a building, or larger building part (e.g., IFC model, PCD). Entity Space has the subclasses Point Space, Area Space (vertical or horizontal) and Volume Space [5]. While these concepts and classifications apply to almost all spatial elements, one in particular holds my special personal interest. "Technical drawings contain information about the complete geometry of the building. In contrast to point clouds, plans also show the dimensions of non-visible components, such as the foundation. Plans are therefore indispensable for extracting the entire building geometry." [6, p. 31]. Their ability to hold extensive geometric and topological data through a set of lines and its widespread use on-site [7] makes them arguably the primary tool for spatial information transfer in today's AECO industry. A spatial relation processor could potentially allow projections of different data types onto technical drawings, prompting the research question:

"How can spatial relations between technical drawings and geometric data be identified through spatial alignment to enable semantic data transfer?"

To subdivide the objective, the research question can be broken down into these sub-questions:

"As what does the virtual geometric representation of a technical drawing classify?"

"How should geometric data be stored in RDF?"

"How can spatial relations be calculated based on RDF data as input?"

To resolve the research questions, the design science research method was employed. By applying its iterative cycles of analysis, development, evaluation and diffusion, a framework was conceptualized, comprising several smaller proof-of-concept-functions for data transformations and geometric operations. These analysis, evaluation and diffusion cycles helped to outline competencies and limitations of the developed system, continually shaping and shifting its functionality to address the research question.

2 Related Work

Çelik et al. conducted a comparative analysis of a hand-defined filter-based (HDFs) and a CycleGAN-based cleanup method for technical drawings. Their findings indicate that the CycleGAN model outperforms HDFs in redrawing lines, geometric shapes, text units, and numbers, thereby enhancing edge detection (e.g., with Sobel filters) and improving OCR accuracy. [8]. Recent works by Faltin et al. present a retrospective modelling approach based on spatially arranged drawing views. Utilizing a Keypoint R-CNN (Region-based Convolutional Neural Network), they were able to simultaneously detect elevation and section symbols. Subsequently, they were able to automatically determine the 2D plans spatial alignment relative to each other [9] [10]. Their work paves the way for projecting geometric data onto technical drawings thorugh respective spatial alignment. The remaining questions on how to store a plans spatial data in a machine readable format and how that data can be queried or linked to other entities is adressed by Zhang and Beetz: BimSPARQL is a framework extending

SPARQL's querying capabilities in regard to building data [11]. "Unlike GeoSPARQL, it is an academic research proposal on incorporating building-related semantics. It introduced domain-specific functions for schema-level semantics, instance-level semantics, and spatial reasoning" [12, p. 8]. Similar to BimSPARQL, Nikitopoulos et al. present the "first parallel and scalable in-memory solution to the problem of spatio-temporal SPARQL query processing" [13, p. 648]. Their proposed DiStRDF System (Distributed spatiotemporal RDF engine) comprises a processing and storage layer. Standard SPARQL queries and separate spatio-temporal constraints are computed by a processing layer which communicates with a storage layer, responsible for distributed storage of RDF data [13]. A Spatial query system utilizing a multi-LLM agent framework, allowing natural language relation and entity queries through triangulated bounding boxes of IFC elements, has been presented by Li et al. [14]. Limited to concentrated geometric data of structural elements (e.g., wall, beam) and general spatial queries, this paper still recommends research in the field of automatic inferencing of missing spatial information in BIM models. Their work furthers the idea of verbally specifying spatial data and queries.

3 Methodology

To analyse and evaluate spatial relations between various building project data types, a three-dimensional virtual test environment was established. Technical drawings were spatially aligned using point cloud data as a topological reference, employing basic transformation commands: location, rotation, and scaling. The 3D geometry of building components depicted in technical drawings was created, resembling the geometric foundation of IFC-elements. Fictional damages were modeled as bounding boxes, representing the result of a conversion process developed by Göbels and Beetz that translates natural language damage reports into geometries [15]. Figure 1 shows the mentioned data-types, superimposed and projected onto a technical drawing. This model was based on a dataset of the Aachen Cathedral, Germany.

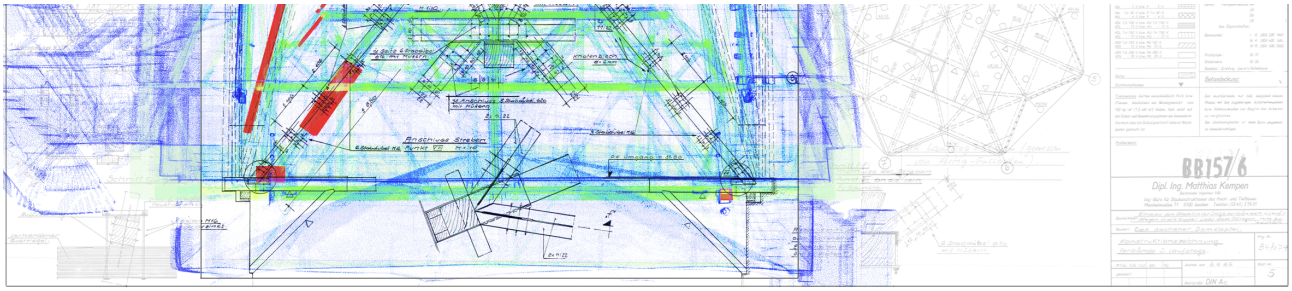


Figure 1: Heterogeneous building data superimposed and projected onto technical drawing. Blue=PCD, green=IFC-elements, red=Damage OBBs.

3.1 Data quality

Despite *limited data heterogeneity* necessitating the creation of additional artificial geometric content to examine potential topological relations between technical drawings and other types of data, the Aachen Cathedral dataset was well-suited for this task due to several factors. It contains a *large number of high-resolution technical drawings* with varying scales, viewing directions, and annotations such as text, measurements, and axes. *Extensive point cloud data* (PCD) provides complete coverage of both the roof's interior and exterior, serving as an accurate spatial reference (asset) for all entities.

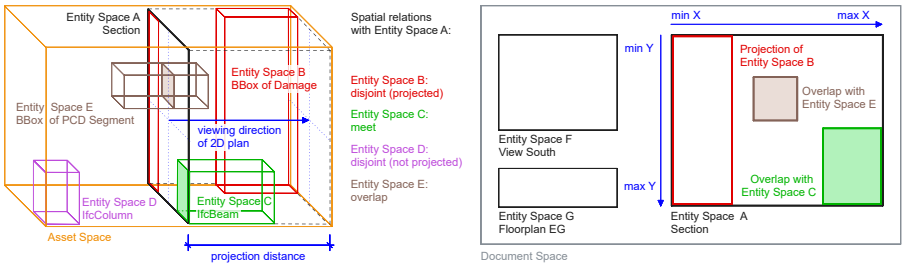


Figure 2: Depiction of technical drawing inside asset and document with spatial relation to multiple entities as well as its projection distance.

The *structural complexity* enables a comprehensive analysis of edge cases in technical drawing orientation and informs a broadly applicable alignment method.

3.2 Accuracy of spatial alignment

The experimental approach highlighted the practicality of spatial alignment of entities with component-specific accuracy to minimize data quality loss during information transfer onto technical drawings. A study on dynamic accuracy levels based on drawing scales and desired information clarity exceeds the scope of this work; therefore, a recommendation for future research is proposed. For now, metric measurements in this study are presented as approximately x.xx m.

3.3 Topological relations

While spatial relations of technical drawings and other data types can be expressed with the vocabulary, presented by Randell et al. [3], only a subset shows relevance when examining the test environment. Technical drawings represent planes in 3D space. Relations such as equal, covers, inside, and contains are rare and typically arise when multiple drawings occupy the same location and share axis alignment. For instance, a vertical section may cover a construction detail of a specific sub-area (coveredBy). Alternatively, two plans may be geometrically and topologically equal by sharing location, axis alignment, and scale while differing in attribute data. The most frequently observed relations are disjoint, meet, and overlap. The relation coveredBy mostly classifies the connection between a technical drawing encapsulated by a bounding box representing a building space. Thus, functions emerging from this work for determining spatial relations should primarily focus on classifying disjoint, meet, and overlap.

3.4 Geometric characteristics of technical drawings

The specification of technical drawings as vertical or horizontal Area Spaces is geometrically correct but topologically insufficient. Some technical drawings do not overlap with all depicted entities, making the calculation of linked elements difficult. A clarification on their spatial extent is needed. Their geometric representation, a plane, in addition to its max projection distance creates an aligned bounding box which encapsulates the plan's spatial influence, rendering the geometric understanding of technical drawings for this work as an Area Space parented to a Volume Space, or just a Volume Space (Fig. 2).

3.5 Conceptualizing a spatial query and processing framework

The following competency questions (CQ) frame the desired functional extent of the spatial query system for a heterogeneous spatial building dataset. CQ1: "Select all bounding boxes of damages

classified as water damages which overlap with the Volume Space defining the spatial influence of Drawing A". CQ2: "Select technical drawings with a local z to global -z axis alignment, overlapping with Entity Spaces attributed with maintenance work". A system capable of adressing these questions was framed (Fig. 3).

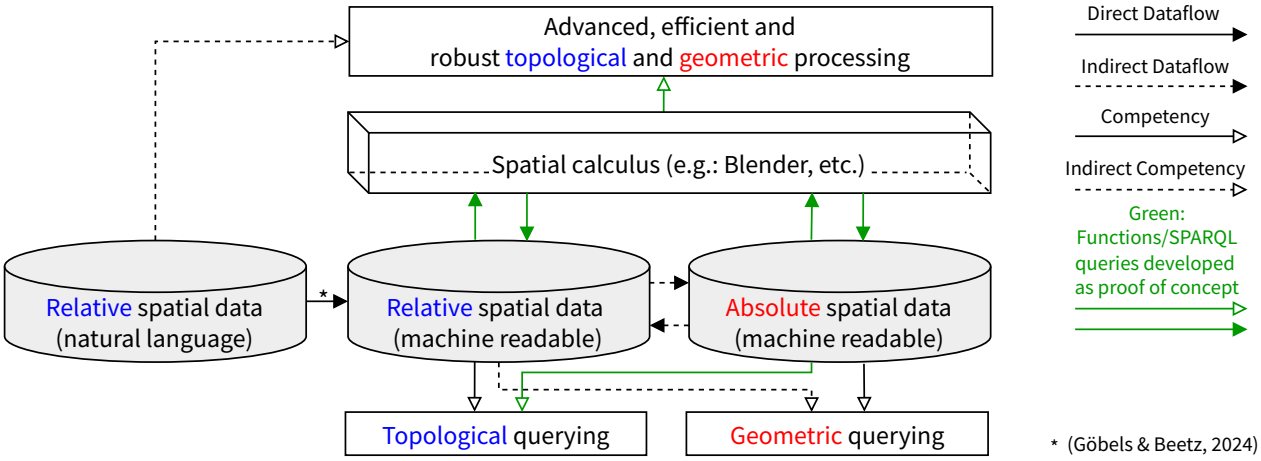


Figure 3: Spatial query and processing Framework

4 Results

The following table 1 shows the results of a comparative analysis of overlap detection algorithms in different environments: (1) Blender python script, (2) SPARQL query.

Table 1: Processing times of overlap detection in milliseconds in a one against all scenario. No.E.S=Number of Entity Spaces, BL=Blender, SP=SPARQL, CE=CenterExtent-BoundingBox algorithm, SP=Sweep and Prune algorithm, MM=MinMax algorithm

No.E.S.	BL.CE.	BL.SP.	SP.CE.	SP.MM.
10	0	0	160	140
55	10	5	209	154
105	19	10	278	174
303	24	30	474	308
507	40	40	645	419

5 Discussion

As what does the virtual geometric representation of a technical drawing classify? For the objective of calculating spatial relations based on the maximum spatial extent, the virtual geometric representation of a technical drawing should be classified as a bounding box or rectangular volumetric space, as specified in section 3.4.

How should geometric data be stored in RDF? While Haverkort [16] outlines the advantage of using simplified representations (bounding boxes) of geometric data, Bonduel et al. [17] demonstrate how linking of multiple geometric representations allows geometric processing on different levels of detail. These approaches sparked the evaluation of different bounding box types regarding the efficiency of detecting overlaps between them using the presented framework. In the Blender python scripts, as

well as the SPARQL queries, the MinMax bounding box (AABB) performed best, as the maximum and minimum values, needed for detecting overlaps are already precomputed. By incorporating axis alignment statements, Entity Spaces can be aligned relative to other entities or Asset Spaces. This bounding box RDF-structure matches with the current working draft by Göbels et al. [5].

How can spatial relations be calculated based on RDF data as input? The comparative analysis' results outlined SPARQL's poor efficiency and accuracy for querying overlaps between Entity Spaces, especially in all against all scenarios. The purposefulness of importing simplified geometric data for complex spatial processing was emphasized. A first rough selection, made by a SPARQL query could filter relevant entities for import. A second more accurate pass inside the spatial calculus could sort out false positives. Their geometric representations could then be used for more complex geometrical or topological processing: Enriching triple-stores with spatial links, transferring semantics based on spatial filters or projecting entities' geometric data onto other Entity- or Document Spaces e.g., technical drawings.

Algorithm 1: Exemplary spatial data export as RDF in compliance with DIN EN ISO 21597-2

```
ex:Section_A-A a ex:AreaSpace ;
        icdd:hasLinkElement ex:Section_A-A_LinkElement .
ex:EntitySpace_1 a ex:VolumeSpace ;
        ex:componentType ex:Beam .
        icdd:hasLinkElement ex:EntitySpace_1_LinkElement .
ex:SpatialLink a icdd:DirectedLink ;
        icdd:hasFromLinkElement ex:EntitySpace_1_LinkElement ;
        icdd:hasToLinkElement ex:Section_A-A_LinkElement .
```

Algorithm 2: SPARQL query retrieving the component type of an Entity Space, depicted on a technical drawing via an ICDD-link

```
SELECT ?plan ?entitySpace ?projectedComponentType
WHERE {
    ?link icdd:hasToLinkElement ?toLinkElement ;
          icdd:hasFromLinkElement ?fromLinkElement .
    ?plan icdd:hasLinkElement ?toLinkElement ;
          rdf:type ex:AreaSpace .
    ?entitySpace icdd:hasLinkElement ?fromLinkElement ;
                 ex:componentType ?projectedComponentType .
}
```

Based on the comparative analysis results and the analysis of spatial structures and characteristics of technical drawings in R^3, a framework of export, import, and processing functions is introduced, transforming data from spatial vocabularies into geometries for advanced processing (insert figure). Similar to DiStRDF [13], the system separates storage and processing. Unlike Apache Spark [18], it imports spatial RDF data into a spatial calculus, supporting diverse geometric and topological operations. Blender was used as an exemplary calculus via its Python API. In theory, the developed functions could be enhanced to handle more complex spatial relation processing and integrated into a

single, software-agnostic function, improving portability. Besides the vendor dependency of the current framework draft, limiting its diffusion or implementation as a web service, the system does not work with oriented bounding boxes, which would enable more accurate geometric processing.

6 Conclusion

This paper presents a system for bridging gaps between relative and absolute location data in hopes of advancing the broader objective of linking semantic data and projecting geometric data through respective spatial alignment. The framework consists of an RDF-to-geometry converter, a spatial relation processor, and a geometry-to-RDF converter (Fig. 3).

Future work regarding comparative analysis of spatial indexing methods could enable the implementation of spatial functions via standard-compliant SPARQL queries without needing an external processing engine. For an advanced diffusion of the system, its implementation as spatial functions, comprising a web service, could allow clients to communicate with a database and a spatial processing engine that returns query results or updates the RDF database based on entity queries and relational queries as input.

Converting an Entity Space's three-dimensional Cartesian coordinates to local technical drawing pixel coordinates or normalized values (for vector-based formats) would allow projecting geometric data onto technical drawings. Developing a viewer to overlay an Entity Space's local pixel coordinates on the corresponding technical drawing could enhance usability and promote sustainable dissemination within the industry.

Acknowledgement This research is part of the Raumlink project funded by the German Research Foundation (DFG) – Project number 501812634.

References

[1] D. Papadias and Y. Theodoridis, "Spatial relations, minimum bounding rectangles, and spatial data structures", *International Journal of Geographical Information Science*, vol. 11, Mar. 1997. DOI: 10.1080/136588197242428

[2] R. Chbeir, Y. Amghar, and A. Flory, "Novel indexing method of relations between salient objects", in *Effective Databases for Text & Document Management*, IGI Global, 2003.

[3] B. Nebel, C. Rich, and W. R. Swartout, *Principles of Knowledge Representation and Reasoning: Proceedings of the Third International Conference (KR '92)*. M. Kaufmann, 1992.

[4] A. Göbels and J. Beetz, "Relative Location Ontology: An ontological model for representing directional topological relationships between spatial entities in oriented space", in *LDAC 2024 – 12th Linked Data in Architecture and Construction Workshop*, Bochum, Germany, 2024.

[5] A. Göbels, O. Schulz, and J. Beetz, "Towards a Common Digital Space: Proposing a Schema for Spatially Linking Heterogeneous Resources", in *Proceedings of the 41st International Conference of CIB W78*, Morocca, Marrakech, 2024.

[6] A. Borrmann et al., "Technologien zur Generierung digitaler Zwillinge als Grundlage für Betrieb und Instandhaltung baulicher Infrastruktur", 2022.

[7] K. A. Bisharat, *Construction Graphics: A Practical Guide to Interpreting Working Drawings*. John Wiley & Sons, Mar. 2025.

[8] F. Çelik, B. Faltin, and M. König, “Improving 2D Construction Plans with Cycle-Consistent Generative Adversarial Networks”, May 2022. DOI: 10.1061/9780784483893.007

[9] B. Faltin, P. Schönfelder, and M. König, “Inferring Interconnections of Construction Drawings for Bridges Using Deep Learning-based Methods”, in *ECPPM 2022 - eWork and eBusiness in Architecture, Engineering and Construction 2022*, CRC Press, 2023.

[10] B. Faltin, P. Schönfelder, D. Gann, and M. König, “Reconstructing as-built beam bridge geometry from construction drawings using deep learning-based symbol pose estimation”, *Advanced Engineering Informatics*, vol. 62, Oct. 2024. DOI: 10.1016/j.aei.2024.102808

[11] C. Zhang, J. Beetz, and B. de Vries, “BimSPARQL: Domain-specific functional SPARQL extensions for querying RDF building data”, *Semantic Web*, vol. 9, Sep. 2018. DOI: 10.3233/SW-180297

[12] D. Elshani et al., “Comparative Analysis of Approaches for Geometric Data Representation in RDF”, 2025.

[13] P. Nikitopoulos, A. Vlachou, C. Doulkeridis, and G. A. Vouros, “Parallel and scalable processing of spatio-temporal RDF queries using Spark”, *GeoInformatica*, vol. 25, Oct. 2021. DOI: 10.1007/s10707-019-00371-0

[14] A. Li, P. K.-Y. Wong, X. Tao, J. Ma, and J. C. P. Cheng, “An interactive system for 3D spatial relationship query by integrating tree-based element indexing and LLM-based agent”, *Advanced Engineering Informatics*, vol. 66, Jul. 2025. DOI: 10.1016/j.aei.2025.103375

[15] A. Göbels, F. Rivadeneyra, and J. Beetz, “Transfer of implicit semi-formal textual location descriptions in three-dimensional model contexts”, in *Proceedings of the 2023 European Conference on Computing in Construction and the 40th International CIB W78 Conference*, 40. International CIB W78 Conference, Heraklion (Greece), 10 Jul 2023 - 12 Jul 2023, Jul. 2023. DOI: 10.35490/EC3.2023.268

[16] H. J. Haverkort, “Results on Geometric Networks and Data Structures.”, May 2004.

[17] M. Bonduel, A. Wagner, P. Pauwels, M. Vergauwen, and R. Klein, “Including widespread geometry formats in semantic graphs using RDF literals”, in *EC3 Conference 2019*, ser. Computing in Construction, vol. 1, European Council on Computing in Construction, 2019. DOI: 10.35490/EC3.2019.166

[18] M. Zaharia et al., “Resilient distributed datasets: A {Fault-Tolerant} abstraction for {In-Memory} cluster computing”, in *9th USENIX Symposium on Networked Systems Design and Implementation (NSDI 12)*, 2012.

Datengetriebener Digital Twin für das Facility Management: Integration von Building Information Modeling, Multi-Modell-Container und Large Language Models

Thujiethaa Thuraisingam[1], Nijanthan Mohan[1] and Rolf Groß[1]
[1]Chair of Institute of Smart Building Engineering, Fachhochschule Aachen, Bayernallee 9, 52066 Aachen
E-mail(s): thujiethaa.thuraisingam@alumni.fh-aachen.de

Abstract: In Zeiten fortschreitender Digitalisierung im Bauwesen gewinnt die Integration heterogener Datenquellen zunehmend an Bedeutung. In diesem Beitrag wird ein Ansatz vorgestellt, der Building Information Modeling (BIM), Multi Model Container (MMC) und Large Language Models (LLMs) kombiniert, um einen datengetriebenen Digital Twin (DT) zur Unterstützung des Facility Managements (FM) zu realisieren. Das vorgestellte Konzept ermöglicht die Verknüpfung von statischen Gebäudedaten aus BIM-Modellen mit dynamisch erfassten Sensordaten innerhalb eines MMC, wodurch ein detailliertes digitales Abbild des Gebäudezustands entsteht.

Zur Validierung wurde ein Büroraum per Laserscan erfasst und mit synthetischen Sensordaten im MMC verknüpft. Ein Python-Skript bereitete die Daten für LLM-Abfragen auf. Erste Tests zeigten vielversprechende, jedoch noch verbesserungswürdige Ergebnisse.

Die Ergebnisse zeigen, dass die Kombination der Technologien die Datenintegration im FM verbessern kann und eine Basis für weitere Forschung bildet.

Keywords: Digital Twin, Facility Management, Building Information Modeling, Multi-Modell-Container, Large Language Models

DOI: 10.18154/RWTH-CONV-254923. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Einleitung

Steigende Anforderungen an Effizienz, Nachhaltigkeit und Transparenz erhöhen im Facility Management (FM) den Bedarf an Echtzeitinformationen. Klassische FM-Systeme stoßen jedoch oft an ihre Grenzen, wenn es darum geht, heterogene Datenquellen, wie BIM-Modelle, Sensordaten und Betriebsdokumente, konsistent zusammenzuführen und effizient nutzbar zu machen. [1], [2] Im Projekt „DiggiTwin“ wurde ein prototypischer Digital Twin (DT) entwickelt [3].Er kombiniert ein Building Information Modeling (BIM)-Modell mit raumklimatischen Sensordaten im Multi-Modell-Container (MMC). Hieraus resultiert ein interoperables digitales Abbild eines ausgewählten Büroraums. Ergänzend

wurde getestet, ob ein generisches Large Language Model (LLM) natürliche Abfragen dieser Daten unterstützen kann. Der vorliegende Beitrag zielt darauf ab, einen prototypischen DT zu konstruieren und das Potenzial sowie die Grenzen eines LLM-basierten Zugriffs auf die zugrunde liegenden Gebäudedaten zu demonstrieren. Als Demonstrationsraum wurde ein Büroraum in einem Hochschulgebäude ausgewählt. Erste Tests zeigen, dass das LLM relevante Informationen bereitstellen kann, abhängig von Datenqualität und Promptgestaltung.

2 Grundlagen

In den Abschnitten 2.1–2.5 werden in kurzer Form die Grundlagen zu BIM, DT, MMC, LLM sowie der FM-Datenbedarf dargelegt. Damit wird der Rahmen für die darauffolgende Umsetzung geschaffen.

2.1 Building Information Modeling

BIM führt geometrische und alphanumerische Daten in einem objektorientierten 3D-Modell zusammen [4], [5]. Der IFC-Standard sichert Interoperabilität. Im Betrieb gewinnt BIM zusätzlichen Wert, da es die statische Basis für DT-Anwendungen im FM bildet. [5]

2.2 Digital Twin

Ammar et al. (2022) identifizieren Monitoring, Anomalieerkennung und prädiktive Instandhaltung als zentrale Einsatzszenarien von DTs im Gebäudebetrieb [6], während Zahedi et al. (2024) in einer Fallstudie zeigen, dass sich damit insbesondere die Energieeffizienz von HVAC-Systemen optimieren lässt [7]. Voraussetzung für solche Anwendungen ist eine Datenstruktur, die heterogene Quellen zuverlässig verknüpft – hier bietet der von Nidhal Al-Sadoon an der Technischen Universität (TU) Dresden entwickelte MMC, bereitgestellt im Tool BIMs-Integrator / MMC ENG, eine geeignete Grundlage [8].

2.3 Multi-Modell-Container

Die Norm DIN 18290-1:2023 beschreibt den Ansatz eines MMC [9]. In dieser werden mehrere Fachmodelle, darunter IFC-Dateien, Leistungsverzeichnisse und Sensordatensätze, gemeinsam mit einem Linkmodell archiviert. Das genannte Modell bildet die Beziehungen auf elementarer Ebene ab. Für diese Arbeit wird die MMC-Implementierung „BIMs-Integrator / MMC ENG“ genutzt, die Nidhal Al-Sadoon an der Technischen Universität Dresden entwickelt und bereitgestellt hat [8]. Das Tool bietet Versionierung, Zugriffskontrolle und eine automatisch auswertbare Verknüpfung der enthaltenen Modelle. [10], [11], [8]

2.4 Large Language Models

LLMs ermöglichen semantische Suche und Dialoganwendungen [12], [13]. Da sie nur Text verarbeiten, müssen Sensordaten oder IFC-Eigenschaften vorab textuell oder tabellarisch aufbereitet werden. Spezialisierte FM-Modelle fehlen bisher, daher werden meist generische Modelle mit Projektdaten angereichert [14].

2.5 Informationsbedarf im Facility Management

Das im Rahmen der Fallstudie durchgeführte Interview mit dem FM des ausgewählten Gebäudes identifizierte drei zentrale Datenbereiche, die im täglichen Betrieb von besonderer Relevanz für die Informationsbereitstellung sind.

Tabelle 1: Ergebnisse des Interviews zum Informationsbedarf im Facility Management

Datenbereich	**Typische Inhalte (Interview)**	**Hauptnutzen für das FM**
Energieverbrauch & Effizienz	Strom- und Wärmeverbrauch, CO_2-Lasten; Vergleich einzelner Gebäudeteile	Identifikation ineffizienter Anlagenteile und Kostenkontrolle
Raum- & Anlagenklima	Temperatur (Räume, Heizkörper), Luftqualität / Feuchte, Wärmefluss (Vor-/Rücklauf)	Regelung von Lüftung und Heizung, frühzeitige Anomalieerkennung
Nutzungs- & Zustandsdaten	Anwesenheitssensoren, Stör- bzw. Fehlermeldungen wichtiger Technik (Heizung, Lüftung, Kälte)	Last- und Flächenmanagement, gezielte Wartungsplanung

Die Befragung ergab, dass die Informationen bislang in voneinander getrennten Systemen abgelegt sind. Entsprechend besteht beim FM das Bedürfnis, sie aus einer integrierten Quelle abrufen zu können, um Betriebsentscheidungen effizienter zu unterstützen.

3 Methodik und Systemarchitektur

In diesem Kapitel wird der Aufbau des DT-Prototyps sowie die technische Pipeline zur automatisierten Datenabfrage über ein LLM beschrieben. Abbildung 1 stellt den Prozess in abstrahierter Form dar und fokussiert auf die wesentlichen Schritte der Datenintegration und LLM-Anbindung. Eine detaillierte Darstellung von Schnittstellen, Datenflüssen und Implementierungsschritten erfolgt nicht, da es sich um einen prototypischen Proof of Concept handelt. Im Prototyp wurden verschiedene LLMs hinsichtlich

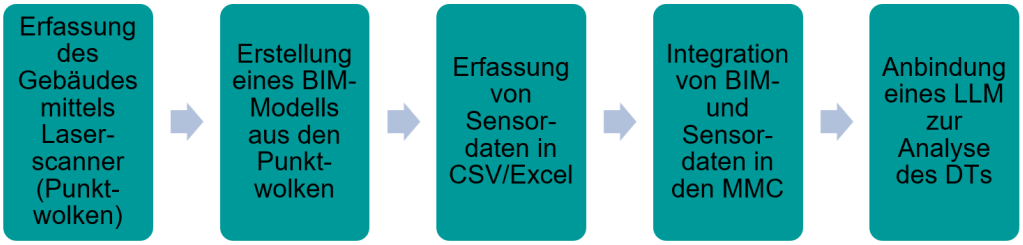


Abbildung 1: Gesamtprozess zur Erstellung des datengetriebenen Digital Twins

ihrer Antwortqualität auf die definierten Anfragen durch manuelle Evaluierung beurteilt. Basierend auf der höchsten erzielten Konsistenz und Präzision wurde das Modell „deepseek-r1:32b“ [15] ausgewählt. Zur Auswahl des Modells wurden zunächst verschiedene Open-Source-LLMs getestet, darunter LLaMA- und Mistral-Modelle vergleichbarer Größe sowie mehrere Varianten der DeepSeek-Familie [16], [17]. Die kleineren Modelle lieferten, unabhängig vom Anbieter, wiederholt unpräzise oder fehlerhafte Antworten. Innerhalb der DeepSeek-Modelle zeigte insbesondere die 32B-Variante im direkten Vergleich die konsistentesten und präzisesten Ergebnisse. Kommerzielle, lizenzpflichtige Modelle wurden in diesem Kontext nicht berücksichtigt, da der Fokus auf frei zugänglichen Open-Source-Modellen lag. Die Tests wurden auf einem leistungsstarken CAD-Arbeitsplatzrechner mit

großem GPU-Speicher (ca. 48 GB VRAM) durchgeführt, um auch größere Modellvarianten verarbeiten zu können.

3.1 Datenbasis und Fallstudie

Zur Veranschaulichung wurde ein Büroraum im Gebäude der Fallstudie als Demonstrationsraum ausgewählt. Abbildung 2 zeigt die Modellansicht des ausgewählten Raums. Die geometrische Erfassung

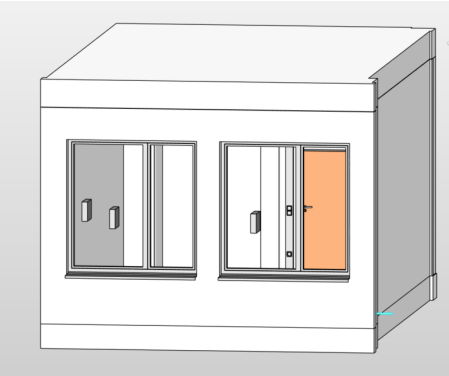


Abbildung 2: Modellansicht des ausgewählten Büroraums im BIM-Modell

wurde mittels Laserscanning durchgeführt, wobei die resultierende Punktwolke in ein IFC-kompatibles BIM-Modell überführt wurde. Im Rahmen der Untersuchung wurden ergänzend synthetische Sensordaten, darunter die Temperatur, die Luftfeuchte sowie der CO_2-Gehalt, im 5-Minuten-Raster generiert. Die Verknüpfung der beiden Datenquellen erfolgte mittels des BIMs-Integrators / MMC ENG in einem MMC. Die Sensorbox wurde mit einer Globally Unique Identifier (GUID)-basierten Verknüpfung zur Zeitreihe ausgestattet. Der Container wurde versioniert und mit Zugriffskontrollen versehen.

3.2 Datenaufbereitung und Antwortpipeline

Zur Beantwortung von Fragen, welche in natürlicher Sprache formuliert wurden, wurde ein mehrstufiges Abfragesystem eingesetzt, welches strukturierte Vorverarbeitung mit generativer KI kombiniert.

Datenextraktion

Die relevanten Informationen aus dem BIM-Modell (beispielsweise Flächen, Materialien und U-Werte) sowie aus der Sensordatei wurden unter Verwendung eines Python-Skripts extrahiert und in tabellarischer Form aufbereitet [18]. Diese Daten bilden die Grundlage für die spätere Analyse und wurden ebenfalls im MMC abgelegt.

Dokumentenzuordnung

Das LLM entscheidet bei jeder Nutzerfrage, auf welches Dokument sich die Anfrage bezieht. Die Entscheidung basiert auf der semantischen Ähnlichkeit zwischen der Frage und der Dokumentenbeschreibung.

Pfadentscheidung

Im darauffolgenden Schritt wird mittels eines Python-Skripts evaluiert, ob die Datenquelle eine ausreichende Struktur aufweist, um eine regelbasierte Analyse zu ermöglichen. Sofern die Frage z. B. eine exakte Zeitangabe enthält, wird ein entsprechendes Skript aktiviert, welches den zugehörigen Messwert errechnet. Bei offenen oder unpräzisen Anfragen erfolgt ein automatischer Wechsel in einen

generativen Pfad, in dem die Antwort durch ein Sprachmodell generiert wird.

Kontextgenerierung und Antwort

Im generativen Pfad wird ein Retrieval-Augmented-Generation-(RAG)-Ansatz eingesetzt. RAG bezeichnet ein Verfahren, bei dem ein Sprachmodell durch die semantische Suche in externen Datenquellen mit kontextrelevanten Informationen angereichert wird [19]. Die abgerufenen Inhalte werden in den Prompt integriert und dienen als Grundlage für die Generierung einer Antwort in natürlicher Sprache. Hierfür werden die Dokumente in kurze Text- oder Tabellenabschnitte zerlegt, mittels Embeddings vektorisiert und in einer semantischen Datenbank abgelegt. Bei einer Nutzeranfrage wird anhand der semantischen Ähnlichkeit eine Auswahl relevanter Abschnitte getroffen, die dem LLM als zusätzlicher Kontext dienen. Das Sprachmodell generiert anschließend eine Antwort in natürlicher Sprache, die auf diesen Kontextinformationen basiert.

Antwortausgabe

Die finalen Antworten enthalten, sofern verfügbar, ergänzend Verweise auf die Datenquelle oder die verwendeten Parameter. Die Ausgabe zeichnet sich dadurch aus, dass sie auch nicht-technischen Anwendern eine Interpretation ermöglicht. Abbildung 3 zeigt exemplarisch eine Konsolenausgabe für eine abgefragte IFC-GUID.

```
Finally, I should present the answer in full sentences as per the user's instruction, ensuring it's clear and concise.
</think>

Die Daten, die zur IFC GUID `0XzHC1b_ D99w3rLG3YlOsV` gehören, sind folgende:

- **Category:** OST_ Walls
- **Name:** bw Flur TÜrrahmen
- **Type:** Wall
- **Id:** 2497645
- **Length:** 4.71 m
- **Width:** 0.40 m
- **Height:** 3.54 m
- **Area:** 14.04 m²
- **Volume:** 5.62 m³
- **U-Value:** 2.70 W/ m²K
- **Materials:** Mauerwerk - Ziegel, Trockenbau - Kalziumsilikatplatte
```

Abbildung 3: Beispielhafte LLM-Antwort zur IFC-GUID

Begründung der Architektur

Die Testläufe ergaben, dass tabellarisch strukturierte Sensordaten, insbesondere mit zeitlich feingranularen Abfragen, von generativen Sprachmodellen allein nicht zuverlässig verarbeitet werden konnten. Das LLM zeigte wiederholt Inkonsistenzen in Bezug auf die korrekte Selektion und Aggregation der Messwerte. Die vorangegangene Skriptanalyse trug signifikant zur Präzisierung bei und gewährleistete reproduzierbare Resultate. Diese Erkenntnis bildete die Grundlage für die Entwicklung der hybriden Abfragestruktur mit regelbasierter Vorverarbeitung.

4 Ergebnisse

Im Zuge der Überprüfung des entwickelten DT-Prototyps wurden exemplarische Abfragen aus dem FM durchgeführt. Die Analyse ergab, dass das System grundsätzlich fähig ist, strukturierte Gebäudedaten und Sensordaten zu verarbeiten und darauf aufbauend passende Antworten in natürlicher Sprache zu

generieren. Bei gezielten Fragen mit Zeitangaben konnten relevante Messwerte automatisiert identifiziert und korrekt zusammengefasst werden. Allgemeiner formulierte Anfragen wurden semantisch passenden Datenquellen zugewiesen und über das Sprachmodell plausibel beantwortet. Insbesondere die vorgelagerte Analyse durch Skripte trug dazu bei, die Genauigkeit bei zeitlich präzisen Fragestellungen deutlich zu erhöhen. Die Tests basieren jedoch auf einem eng abgegrenzten Prototyp mit beschränkter Datenbasis und einem begrenzten Fragespektrum. Aussagen zur Skalierbarkeit, Robustheit oder Nutzerinteraktion im Regelbetrieb lassen sich daraus noch nicht ableiten. Dennoch zeigen die Ergebnisse, dass die hybride Architektur, die Kombination aus regelbasierter Vorverarbeitung und LLM-basierter Generierung, grundsätzlich geeignet ist, relevante Informationen aus einem DT bereitzustellen.

5 Chancen und Limitationen

Die Fallstudie zeigt, dass die entwickelte DT-Architektur grundsätzlich geeignet ist, heterogene Gebäudedaten automatisiert zu verarbeiten und für das FM nutzbar zu machen. Der niedrigschwellige Zugang über natürliche Sprache reduziert den Schulungsaufwand und ermöglicht auch nicht-technischen Nutzergruppen eine Interaktion mit komplexen Informationen. Die Verknüpfung von Revit-Export und Sensordaten im MMC ermöglicht eine kontextbezogene Zusammenführung baulicher und betrieblicher Merkmale [20]. Dadurch wird beispielsweise die Bewertung raumklimatischer Auffälligkeiten erleichtert. Aufgrund der modularen Architektur lässt sich das System zudem flexibel erweitern, etwa durch die Anbindung zusätzlicher Datenquellen oder Analysepfade und perspektivisch auf weitere Gebäudetypen und Nutzungskontexte übertragen. Gleichzeitig werden Einschränkungen sichtbar, die einer breiten Übertragung auf reale Anwendungsfälle derzeit noch entgegenstehen. Der Prototyp beschränkt sich auf einen einzelnen Büroraum mit synthetischen Sensordaten, sodass Fragen der Skalierbarkeit, Performance und Robustheit bei größeren Gebäudemodellen offenbleiben. Darüber hinaus zeigt sich eine hohe Abhängigkeit der Antwortqualität von der Struktur und Kommentierung der Eingangsdaten, insbesondere uneinheitlich benannte Spalten in CSV-Dateien erschweren die Interpretation. Auch reagierte das System empfindlich auf geringfügige Variationen in der Formulierung von Nutzeranfragen, wodurch es bei strukturarmen Zeitreihen zu fehlerhaften Ergebnissen führen konnte. Die semantische Dokumentenzuordnung hingegen erwies sich als stabil. Insgesamt unterstreicht die Studie das Potenzial LLM-basierter DTs im FM, zeigt jedoch zugleich die Notwendigkeit strukturierter Daten und robuster Modelle für eine zuverlässige Anwendung in der Praxis.

6 Fazit

Die vorliegende Arbeit kommt zu dem Schluss, dass die Kombination aus BIM, MMC und LLMs einen innovativen Ansatz für die datengetriebene Unterstützung im FM darstellt. Im Rahmen einer prototypischen Fallstudie konnte nachgewiesen werden, dass sich BIM-Modelle und Sensordaten in strukturierter Form integrieren lassen und für natürlichsprachliche Abfragen nutzbar gemacht werden können. Der entwickelte DT eines Büroraums fungiert dabei als zentrale Informationsquelle, die heterogene Datenquellen konsistent zusammenführt und für operative Entscheidungsprozesse aufbereitet. Besonders hervorzuheben ist die hybride Architektur der Abfragepipeline, welche regelbasierte Analysepfade mit generativer KI kombiniert. Diese Lösung ermöglichte eine robuste, nachvollziehbare

Beantwortung technischer Nutzeranfragen. Die Resultate des Funktionstests belegen, dass sowohl die Datenklassifikation als auch die Antwortgenerierung in einer Vielzahl realistischer Szenarien zuverlässig funktionierten. Gleichzeitig wurden Limitationen identifiziert, etwa die Sensitivität gegenüber Prompt-Formulierungen, die Abhängigkeit von der Datenqualität sowie offene Fragen zur Skalierbarkeit bei größeren Gebäuden oder umfangreicheren Datensätzen. Aufgrund der Ergebnisse dieser Arbeit lassen sich weitere Untersuchungen ableiten, die darauf abzielen, das entwickelte Konzept breiter in realen Gebäudebeständen anzuwenden. Zukünftig sollte insbesondere die Einbindung weiterer Datenquellen, der Einsatz effizienterer KI-Modelle sowie eine fortlaufende Optimierung der Datenaufbereitung näher untersucht werden. Darüber hinaus eröffnet die Anbindung eines LLM-basierten Systems neue Möglichkeiten für eine intuitive und nutzerzentrierte Schnittstelle im FM, ein Schritt hin zu einem smarteren und klimaneutralen Gebäudebetrieb.

Danksagung

Ein besonderer Dank gilt dem Institut für Smart Building Engineering – dem Team des Projekts DiggiTwin an der FH Aachen, gefördert durch das Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen unter dem Förderkennzeichen 005-2302-0018, Deutschland, an der FH Aachen, für ihren kooperativen Geist, der die Tiefe unserer Arbeit bereichert hat.

Literatur

[1] N. Mohan, R. Groß, K. Menzel und F. Theis, »Opportunities and Challenges in the Implementation of Building Information Modelling for Prefabrication of Heating, Ventilation and Air Conditioning Systems in Small and Medium-Sized Contracting Companies in Germany – A Case Study«, in *Building Information Modelling (BIM) in Design, Construction and Operations IV*, 2021. DOI: 10.2495/BIM210101

[2] N. Mohan, G. Wolf, M. Wogan, J. Beilfuß und R. Groß, »Auf dem Weg zu einem ganzheitlichen Lebenszyklusmanagement von Bauwerken für nachhaltiges Bauen«, in *34th Forum Bauinformatik / 34. Forum Bauinformatik (Bochum, 06. - 08.09.2023)*, Ser. Vorschlag einer Methodik zur digitalen Transformation durch datenbankbasierte Informationsmodelle, 2023. DOI: 10.13154/294-10090

[3] R. Groß, B. Döring, T. Frauenrath und L. Altherr, *Vorhabensbeschreibung - DiggiTwin*, Aachen, 2023.

[4] B. für Verkehr und digitale Infrastruktur - BMVI, *Stufenplan Digitales Planen und Bauen – Einführung moderner, IT-gestützter Prozesse und Technologien bei Planung, Bau und Betrieb von Bauwerken*, Berlin, 2015.

[5] buildingSMART Deutschland, *Vielfalt von Open-BIM*, https://www.buildingsmart.de/sites/default/files/inline-files/Multimediabrosch%C3%BCre_buildingSMART%20Deutschland%20e.%20V_offline_0.pdf, 2024.

[6] A. Ashtarout u. a., »Digital Twins in the Construction Industry: A Perspective of Practitioners and Building Authority«, *Front. Built Environ.*, Jg. 8, S. 834 671, Juni 2022.

[7] F. Zahedi, H. Alavi, J. M. Sardroud und H. Dang, »Digital Twins in the Sustainable Construction Industry«, *MDPI*, 2024.

[8] N. Al-Sadoon, R. J. Scherer und K. Menzel, »Dynamic Multi-model Container Framework for Cloud-Based Distributed Digital Twins«, 2025.

[9] DIN Deutsches Institut für Normung e.V., »DIN 18290-1: Verlinkter BIM Datenaustausch von Bauwerksinformationsmodellen mit weiteren Fachmodellen - Teil 1: Verlinkter Datenaustausch mehrerer Fachmodelle beim Building Information Modeling (Multimodell-Container)«, DIN Deutsches Institut für Normung e.V., Berlin, Techn. Ber., Nov. 2023.

[10] N. Al-Sadoon, R. J. Scherer und K. Menzel, »FROM STATIC TO DYNAMIC INFORMATION CONTAINERS«, in *International Council for Research and Innovation in Building and Construction*, Griechenland, 2023.

[11] N. Al-Sadoon, K. Menzel und R. J. Scherer, *Multi-model Framework for Digital Twin Empowerment*, Sep. 2023.

[12] C. F. Gethmann u. a., *Künstliche Intelligenz in der Forschung*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2022.

[13] M. Seemann, *Künstliche Intelligenz, Large Language Models, Chat-GPT und die Arbeitswelt der Zukunft*. Düsseldorf: Hans-Böckler-Stiftung, 2023.

[14] D. Krütli und T. Hanne, »Augmenting LLMs to Securely Retrieve Information for Construction and Facility Management«, *Information*, Jg. 16, Nr. 2, S. 76, 2025. Adresse: https://www.mdpi.com/2078-2489/16/2/76

[15] D. Inc., *DeepSeek-R1 Release*, https://api-docs.deepseek.com/news/news250120, 2025.

[16] M. AI, *LLaMA 3 Model Card and Prompt Formats*, https://www.llama.com/docs/model-cards-and-prompt-formats/llama4/, 2024.

[17] M. AI, *Mistral Models Overview*, https://docs.mistral.ai/getting-started/models/models_overview/, 2023.

[18] P. S. Foundation, *Python 3.13.2 documentation*, https://docs.python.org/3/index.html, 2025.

[19] P. Lewis u. a., »Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks«, in *Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

[20] A. Inc., *BIM-Werkzeuge für Architekten*, https://www.autodesk.com/de/products/revit/architecture, 2025.

Matrix-Based Approach to Optimizing Disassembly Sequence Planning for Aluminum-Framed Windows

Begüm Aktas[1]
[1]Technische Universität Berlin, Department of Civil Systems Engineering, Berlin, Germany
E-mail(s): b.aktas@tu-berlin.de

Abstract: The construction industry is under pressure to minimize its environmental impact, especially in environment where resources are limited and ecological concerns are increasing. Disassembly has emerged as a crucial strategy for facilitating material recovery at the end-of-life (EoL) stage. However, the disassembly process remains complex and necessitates strategic planning to optimize material recovery and reduce waste. Unlike most products, windows are larger, more intricate, and composed of multiple materials, layers, and subsystems. This complexity requires comprehensive disassembly planning. Graph-based planning often struggles to accurately capture the extensive range of precedence relations present in complex, multi-material products. To address this limitation, matrix-based representations have been introduced to convert graphical data into computable formats. Matrices can effectively represent geometric, material, and technical precedence relations embedded in 2D cross-section drawings, allowing for the transformation of this information into matrix data that algorithms can use to generate disassembly sequences. This study reviews both adjacency and incidence matrices, employing unweighted and weighted data to develop disassembly sequences. It emphasizes the capabilities of these models in capturing the geometric and material-based relations among aluminum-framed window parts within the parametric modeling environment. Additionally, it identifies key factors, dependencies, and challenges associated with the matrix-based disassembly of multi-material window systems.

Keywords: Disassembly Sequence Planning, Matrix, Parametric Design, Material Recovery, Window

DOI: 10.18154/RWTH-CONV-254917. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

The global shift towards resource and material recovery is becoming increasingly important in the built environment, as it helps reduce the environmental impact of the construction industry. By enabling better material recovery through disassembly planning, material re-circulation opportunities alone contribute 60 percent to the CO2 emissions of industrial production within the EU [1]. EoL can ultimately be defined as the inability for the system to fulfill its design function or meet new requirements [2]. Disassembly sequence planning, which aims to determine the optimal disassembly

sequence, is necessary to enhance efficiency and minimize costs [3]. Considering disassembly constraints is important not only at the end of a product's life but also throughout its life cycle, in order to reduce problems related to its exploitation and maintenance [4]. By identifying the steps to improve disassembly and recovery, this study contributes directly to achieving sustainability goals, such as reducing landfill waste, resource efficiency at the EoL, and reducing carbon footprints.

Integration of Circular Economy (CE) strategies demonstrates how material flow analysis can map material value chains and support sustainable building processes, becoming prominent through the integration of disassembly planning. Aluminum-framed windows, due to their widespread use and high material value, play a pivotal role in circularity through material recovery strategies, such as reuse and recycling. The average lifespan of aluminum windows is approximately 20-25 years, but this can vary depending on the type of window, the level of maintenance, and the local climate. Thanks to its long lifespan, high level of recyclability, and low maintenance requirements, aluminum is an extremely suitable material for durable products in the construction industry. Aluminum-framed windows with their high material recyclability offer a significant opportunity to increase material recovery and conserve resources. This aligns with the broader goal of mitigating resource scarcity and environmental degradation. Additionally, improved recovery practices can reduce costs for manufacturers by reintroducing materials into the production cycle, creating closed-loop systems.

2 Point of Departure

Disassembly plays a key role in the CE, particularly for high-value building products such as windows, during demolition, deconstruction, and the dismantling process. The disassembly process is a progressive dismantling of the core into its assemblies and sub-assemblies through a sequence of tasks [5]. The geometry of the parts needs to be considered when disassembling, as CAD files can provide valuable disassembly knowledge by accessing geometric design information [6]. CAD files as 2D cross-section drawings essentially describe the geometry of each window part profile, their relative position and fit (adjacency), and assembly logic, which indicates who touches whom. Geometric feasibility means it is possible to disassemble two sub-assemblies or parts without collision [6]. Generating geometrically feasible disassembly sequences from non-destructive operations requires collision-free precedence relationships. Therefore, the key aspects of geometric disassembly sequence reasoning are to determine the precedence disassembly relationships correctly and in their entirety [6].

The disassembly process is a complex task that necessitates not only topological precedence relationships but also semantic and technical relationships, as well as material-sensitive relations. These relationships are crucial in generating a feasible and robust disassembly sequence planning. The search for precedence relations is indispensable for proper modeling, because those relations have to be known before establishing the set of basic constraints that must be met by the feasible sub-assemblies [7]. The key aspect of these relationships is their ability to be captured in matrix and graph data forms, which then serve as input for mathematical tools and methods. Matrix-based methods, such as those used in computer handling of disassembly constraint relationships, provide a more convenient means for the computer to handle these complex relationships. To ensure that parts can be disassembled without interfering with each other, collision-free precedence relations can be captured and generated in the Adjacency and Incidence Matrix data format. This evaluation can be carried

out using collision analysis on a 3D assembled window model, where collisions between parts are computed based on topological, spatial relationships, and distances between parts. To streamline the collision test between window parts, the frame profile geometries are converted into bounding boxes, which are used to construct relationships in matrix data format for both adjacency and incidence matrices. Bounding box simplification is used to conduct preliminary collision analysis. When bounding boxes do overlap or touch, the components are considered adjacent or colliding [6]. The adjacency matrix, a square matrix (n x n) that captures direct part-to-part relations as node-to-node connections in this study, in binary data format, is primarily used for topological reasoning. If node i is adjacent to node j, the matrix entry AM(i,j) is 1; otherwise, it is 0. On the other hand, the incidence matrix plays a significant role in algorithms for path finding, optimization, and connectivity analysis. It captures the node-to-edge relations as relationships between parts (nodes) and connections (edges). If a part i is involved in a connection j, the matrix entry IM(i,j) is 1; otherwise, it is 0. Therefore, in this study, we adapted Dijkstra's shortest path algorithm to determine the disassembly sequence of windows by traveling node-to-node or node-to-edge paths in a parametric modeling environment that enables the generation of variations of one cross-section in a short time.

3 Methodology

The present study employs a parametric design environment to convert a 2D cross-sectional drawing of a window into a 3D model. Relationships between window parts are identified through collision analysis, using geometric relations and matrix-based data processing. To explore the role of graph-based representations in developing disassembly sequences, the methodology begins with generating 3D window models from 2D drawings. These models are used to determine geometry-based precedence relations among parts. Subsequently, graph-based data are represented in matrix form, with particular attention to adjacency and incidence matrices. Material and interface types are then assigned to each parts to ensure the approach accounts for material characteristics and connection types. Weight computation follows a structural complexity assessment method well-established in the literature [8], [9]. For materials, weights range from 0 (low recyclability) to 1 (high recyclability). Aluminum is assigned a weight of 1.0, while the thermal break is assigned 0.1, reflecting their respective recyclability. Interface types are weighted based on four factors: Connection Type (CT), Connection Accessibility (CA), Independency (ID), and Geometry of the Product Edge (GPE). For instance, in this study, the aluminum-aluminum (mechanical) interface uses screw connections (CT=0.8, CA=1.0), resulting in an average weight of 0.9. The aluminum-thermal break interface averages ID=0.4 (incidental intersection) and GPE=0.4 (overlap). The thermal break-thermal break interface averages CA=0.1 (not accessible), ID=0.1 (full integration), and GPE=0.1 (closed), resulting in a final weight of 0.1.

Once material and interface weights were assigned, we prefer a heuristic-hybrid algorithm combining graph traversal techniques with a directional binary connection matrix. This matrix, generated from the as-parsed part relationships, was used to construct a shortest-path tree from a starting node (representing a window part) via dependency-first traversal. The algorithm prioritizes higher index neighbors and selects the nearest unvisited neighbor at each step, thereby forming a single continuous disassembly path that visits all nodes exactly once. This process is analogous to the Traveling Salesman Problem but adapted for deterministic, material-sensitive disassembly sequencing. We

adapted Dijkstra's shortest-path algorithm into a greedy variant, which selects the locally optimal next step while minimizing total travel distance and cost. The algorithm avoids branching and multiple paths, ensuring a complete and linear disassembly sequence. Material and interface priorities are incorporated through weighted adjacency and incidence matrices, enabling sensitivity to recyclability and connection complexity. The role of graph-based representation, particularly adjacency and incidence matrices, was examined under both weighted and unweighted conditions. The proposed parametric approach was tested on 3D window models of varying complexity (8, 16, 20 parts). This stepwise methodology allowed for direct comparison of disassembly performance across scales and material heterogeneity. Four data configurations were evaluated: unweighted adjacency, weighted adjacency, unweighted incidence, and weighted incidence matrices. Initial validation was performed on three window models, followed by application to increasingly complex assemblies. We plan to extend the analysis to ten window configurations with diverse cross-section details, part counts, and window types, including single-vent and fixed models. The approach is implemented within the Rhinoceros-Grasshopper parametric design environment.

4 Results

We applied adjacency and incidence matrices to evaluate the role of matrix type in disassembly sequence planning under unweighted and weighted traversal schemes to an 8-part assembled 3D window model. Two material combinations are tested: homogeneous (all aluminum) and heterogeneous (aluminum-thermal break combination) within weighted and unweighted adjacency and incidence matrices. Minor differences are observed between adjacency and incidence matrix inputs to Dijkstra's Greedy Traversal algorithm; however, weighting based on material and interface types notably altered disassembly sequences in mixed-material assemblies. In the heterogeneous case, the weighted matrix differed from the unweighted one, prioritizing aluminum material and aluminum-aluminum connections due to stronger separation weights rather than lower thermal break-thermal break interfaces. On the other hand, in the homogeneous case where all parts were aluminum, weighted and unweighted paths overlapped each other, as all weights were uniform (0.9), as shown in Figure 1.

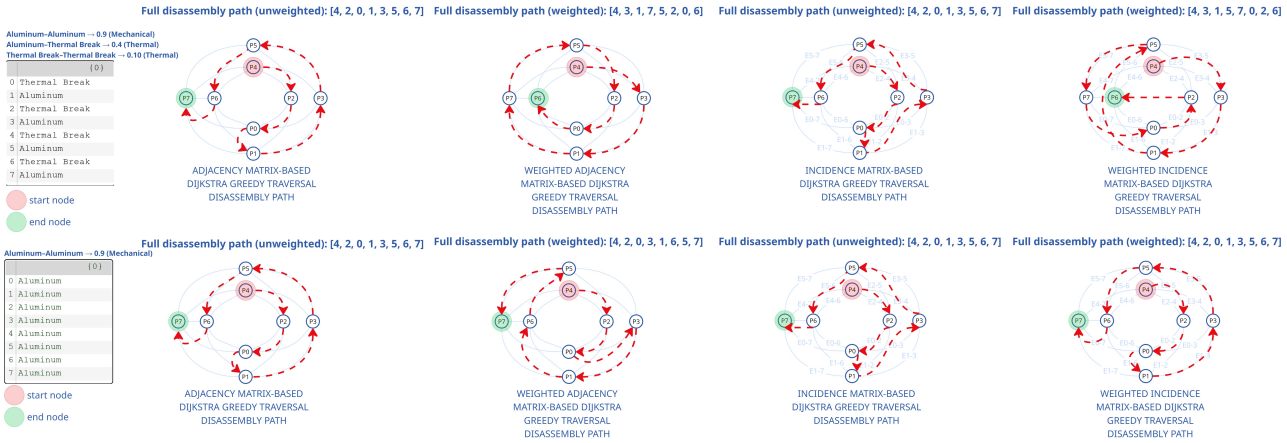


Figure 1: Disassembly sequences of adjacency and incidence within unweighted and weighted matrices for eight-part window

After evaluating the 8-part window, we applied the proposed framework to a 16-part window model, which increased material diversity by incorporating more thermal break parts. More complex disassembly sequence plans are observed, particularly within weighted adjacency and incidence matrices, which tend to produce more notable results. Due to varied material and interface types (aluminum-aluminum (0.9), aluminum-thermal break (0.4), and thermal break-thermal break (0.10)), Dijkstra's Greedy Traversal algorithm generated disassembly sequences more aggressively based on interface weights in these weighted matrices (Figure 2). Dissimilarity between adjacency and incidence matrix-based disassembly sequence development begins to emerge in terms of traversal efficiency and order.

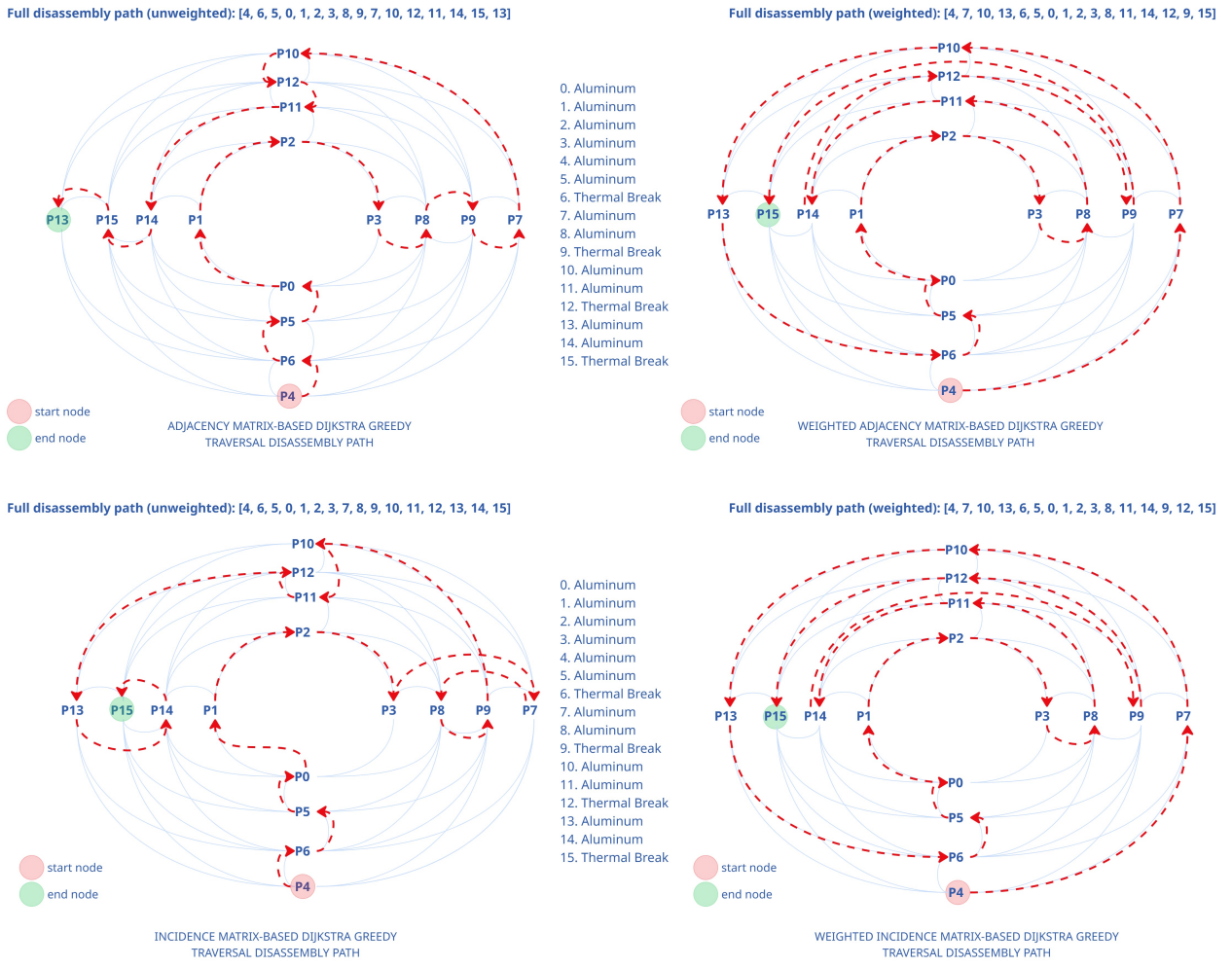


Figure 2: Disassembly sequences of adjacency and incidence under unweighted and weighted matrices for sixteen-part window

We, again, defined the start with 'node 4' for each of these disassembly sequence developments. As expected, we observed that unweighted disassembly sequences of the adjacency matrix-based approach [4, 6, 5, 0, 1, 2, 3, 8, 9, 7, 10, 12, 11, 14, 15, 13] and the incidence-based approach [4, 7, 10, 13, 6, 5, 0, 1, 2, 3, 8, 11, 14, 12, 9, 15] followed a topological/geometric pattern without sensitivity to material and interface types. Small differences between the two matrix types arose from the traversal order due to node-to-node and node-to-edge relations when nodes have equal distances in the graph. From the weighted matrices point of view, a weighted adjacency matrix-based sequence is generated

as [4, 6, 5, 0, 1, 2, 3, 7, 8, 9, 10, 11, 12, 13, 14, 15] and a weighted incidence matrix-based sequence as [4, 7, 10, 13, 6, 5, 0, 1, 2, 3, 8, 11, 14, 9, 12, 15]. These sequences reflect the role of material and interface types in prioritizing strong (aluminum-aluminum) connections, delaying lower weights. This prioritization demonstrates that Dijksta's Greedy Traversal algorithm prioritizes material and interface types, minimizing disassembly effort by postponing weaker or mixed joints. As a result, weighted matrices provide logically better and more comprehensive disassembly sequences.

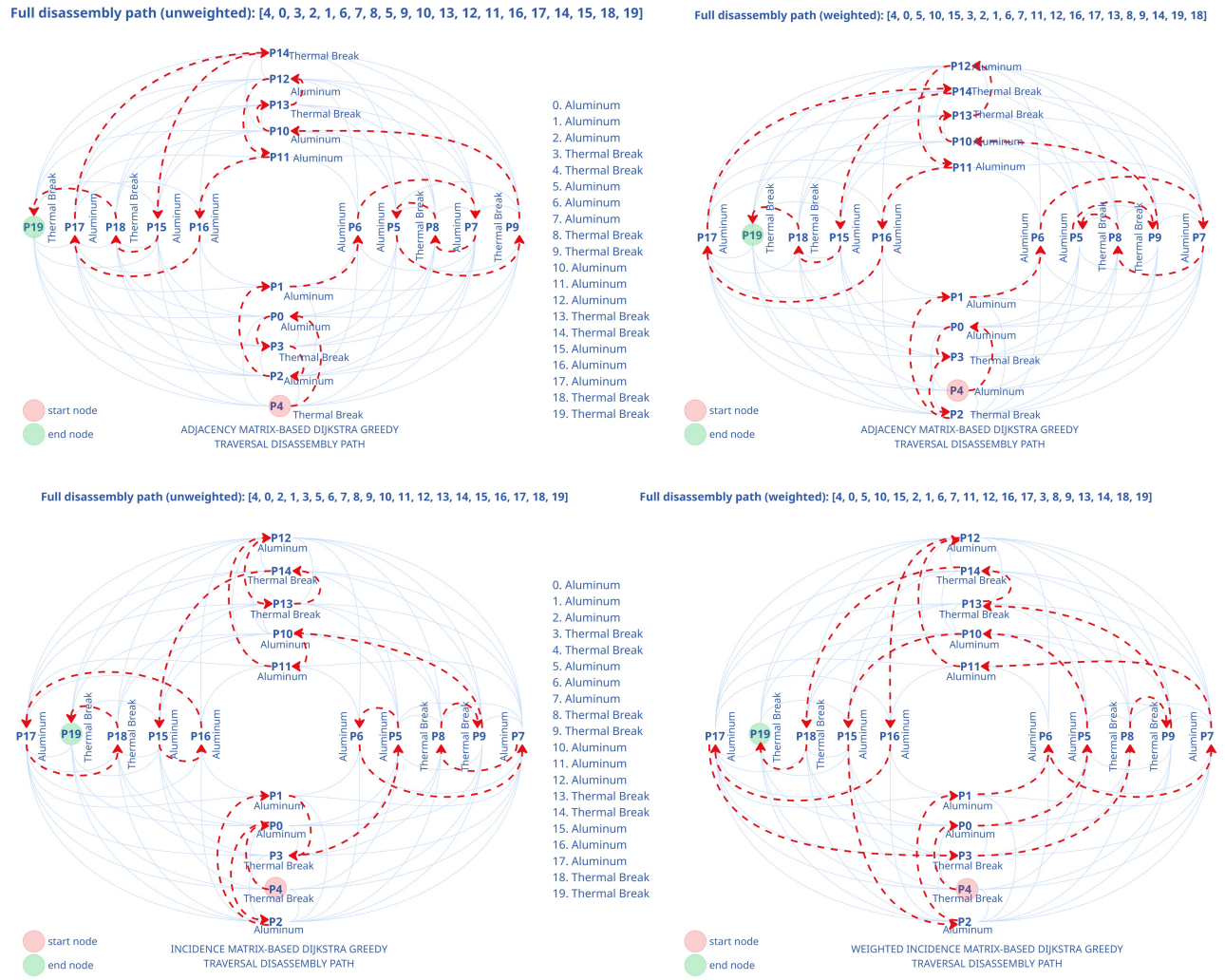


Figure 3: Disassembly sequences of adjacency and incidence under unweighted and weighted matrices for twenty-part window

Once and for all, we extended our window model part numbers to 20 and evaluated the proposed approach under unweighted and weighted adjacency and incidence matrices. Similarly, the 8-part window and 16-part window, as well as unweighted matrices, generate topological disassembly sequences that are insensitive to material and interface types. Nodes, which are window parts, are visited in a generally sequential or radial order based on geometric relations. However, weighted matrices prioritized material and interface type relations, assigning higher weights to aluminum and aluminum-aluminum interfaces, followed by weaker weights. Nodes like 0, 1, 2, 6, 7, 11, and 12 appeared on the disassembly sequence list earlier due to stronger (aluminum-aluminum) relations

(Figure 3). Nodes with more thermal break material and interface types appeared later. Herein, the incidence matrix generates similar but not identical weighted sequences compared to the adjacency matrix. This showed that incidence matrices are more sensitive to edge-based characteristics than adjacency matrices.

5 Discussion

To evaluate different disassembly sequence planning strategies, simulations were conducted on aluminum-thermal break windows with 8, 16, and 29 parts. Using a greedy adaptation of Dijkstra's algorithm, both adjacency and incidence matrices were tested under unweighted and weighted conditions. Material-based weights modeled separation complexity, enabling realistic sequence planning. Unweighted matrices produced similar sequences regardless of material type, reflecting purely geometric and topological relationships. Adjacency matrices emphasized part-to-part relations, while incidence matrices captured more part-to-joint relations, offering finer control over traversal paths. The latter proved more sensitive to shared interfaces, an advantage in complex assemblies where multiple parts meet. Weighted matrices incorporated material-interface logic, prioritizing disassembly according to connection strength ($0.9 > 0.4 > 0.1$). Stronger connections, such as aluminum-aluminum, were removed first, while weaker ones, such as aluminum-thermal break or thermal break-thermal-break, were delayed. This behavior confirms the method's sensitivity to both structural arrangement and material-specific interface properties, demonstrating that Dijkstra's algorithm effectively uses weight data to minimize disassembly effort.

6 Conclusion

This study presents a graph-based disassembly sequence planning approach for aluminum-framed windows, utilizing Dijkstra's Greedy Traversal algorithm within adjacency and incidence matrices under both unweighted and weighted conditions. The results show that while unweighted matrices produce disassembly sequences primarily governed by geometry and topological relations, weighted matrices generate material-sensitive sequences, allowing the algorithm to compute not just geometrically optimal but also material- and effort-aware disassembly sequences. Another key finding is that incidence matrices yield more edge-sensitive results for controlling joint-level disassembly relations than adjacency matrix results. The use of weighted matrices results in more realistic disassembly sequences, aligning with practical expectations for developing disassembly sequences for multilayer and multi-material windows. The proposed framework was validated for robustness and scalability, even for windows with increasing complexity, with 8, 16, and 20 parts. In conclusion, this study shows the practical implications of using graph theory-based models within (un-)weighted traversal algorithms for disassembly sequence planning in parametric design contexts. The proposed framework not only enables a more material- and geometry-informed disassembly sequence planning process but also supports material recovery strategies, making it highly applicable in real-world manufacturing scenarios.

Acknowledgements

We gratefully acknowledge support from the European Union's Horizon Europe research and innovation program under grant agreement No. 101056773 for funding this work under the Reincarnate project.

References

[1] F. Kullmann, P. Markewitz, D. Stolten, and M. Robinius, “Combining the worlds of energy systems and material flow analysis: A review”, *Energy, Sustainability and Society*, vol. 11, no. 1, p. 13, 2021.

[2] R. Hartwell and M. Overend, “Unlocking the re-use potential of glass façade systems”, *Glass Performance Days*, vol. 2019, pp. 273–280, 2019.

[3] Z. Zhou et al., “Disassembly sequence planning: Recent developments and future trends”, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 233, no. 5, pp. 1450–1471, 2019.

[4] M. Kheder, M. Trigui, and N. Aifaoui, “Optimization of disassembly sequence planning for preventive maintenance”, *The International Journal of Advanced Manufacturing Technology*, vol. 90, pp. 1337–1349, 2017.

[5] D. Li, K. Gao, Y. Ren, R. Zhang, and Y. Fu, *Integrating meta-heuristics and a Sarsa algorithm for disassembly scheduling problems with cycle time and hazard coefficients*. OAE Publishing Inc., 2024.

[6] J. P. J. Prioli, H. M. Alrufaifi, and J. L. Rickli, “Disassembly assessment from cad-based collision evaluation for sequence planning”, *Robotics and Computer-Integrated Manufacturing*, vol. 78, p. 102 416, 2022.

[7] A. J. Lambert, “Disassembly sequencing: A survey”, *International journal of production research*, vol. 41, no. 16, pp. 3721–3759, 2003.

[8] K. Sinha and O. L. de Weck, “Empirical validation of structural complexity metric and complexity management for engineering systems”, *Systems Engineering*, vol. 19, no. 3, pp. 193–206, 2016.

[9] B. Alkan, D. Vera, B. Ahmad, and R. Harrison, “A method to assess assembly complexity of industrial products in early design phase”, *IEEE Access*, vol. 6, pp. 989–999, 2017.

Regelbasierte Modellprüfung im Krankenhausbau: Entwicklung eines Regelsatzes zur automatisierten Qualitätskontrolle

Sabine Hartmann[1], Linus Kloesgen[1], Hendrik Morgenstern[1] and Katharina Klemt-Albert[1]
[1]Institut für Baumanagement, Digitales Bauen und Robotik im Bauwesen (ICoM), RWTH Aachen University, Aachen, Germany
E-mail(s): hartmann@icom.rwth-aachen.de

Abstract: Die zunehmende Komplexität moderner Bauprojekte, insbesondere im technisch anspruchsvollen Krankenhausbau, stellt Architekten und Planer vor große Herausforderungen. Als Beispiel zu nennen sind die hohen Anforderungen an den Brandschutz und die damit verbundenen Schnittstellen, welche die Anfälligkeit für Planungsfehler erhöhen und zu kostspieligen Nachbesserungen führen können. Automatisierte modellbasierte Prüfungen von digitalen Gebäudemodellen können dazu beitragen, Planungsfehler frühzeitig zu erkennen, um somit Kosten zu reduzieren sowie die Effizienz im Planungsprozess zu steigern. In diesem Beitrag wird die Erstellung eines Regelsatzes zur automatisierten Qualitätssicherung von Brandschutzanforderungen für den Krankenhausbau vorgestellt. Grundlage bilden die brandschutzrelevanten Anforderungen der Muster-Krankenhausbauverordnung (KhBauVO), implementiert und ausgeführt mit der Solibri Model Checker Software. Es wird aufgezeigt, dass der Regelsatz zahlreiche Verstöße gegen baurechtliche Vorgaben aufdecken kann – etwa unzureichende Treppenmaße oder fehlende Feuerwiderstandseigenschaften bei Türen. Die Ergebnisse verdeutlichen gleichzeitig die Herausforderungen in der Vielfalt landesspezifischer Vorschriften und im hohen Aufwand für die Entwicklung der Regeln und begrenzen die unmittelbare Einsetzbarkeit eines einheitlichen Regelsatzes. Dennoch lässt sich feststellen, dass ein standardisiertes, BIM-gestütztes Prüfverfahren die Planungsqualität verbessern und insbesondere im Baugenehmigungsprozess erhebliche Effizienzsteigerungen erzielen kann. Zukünftige Automatisierungen in der Modelldatenaufbereitung sowie die Integration in behördliche Abläufe könnten die Hürden weiter senken und die Vorteile der regelbasierten Modellprüfung von komplexen Bauwerksanforderungen voll ausschöpfen.

Keywords: Building Information Modeling (BIM), Model Checking, Krankenhausbau, Brandschutz, Automated Compliance Checking (ACC)

DOI: 10.18154/RWTH-CONV-254918. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Einleitung

Die Bauindustrie befindet sich gegenwärtig in einem Prozess des digitalen Wandels. Dabei spielt die Methode des Building Information Modelings (BIM) eine wesentliche Rolle, da sie erhebliche Effizienzsteigerungen in verschiedenen Lebenszyklusphasen eines Bauwerks ermöglicht. Besonders in komplexen Gebäuden mit zahlreichen Schnittstellen besteht ein großes Potenzial zur Optimierung von Planung, Ausführung und Betrieb. Als hochkomplexe Gebäude gelten z. B. Krankenhäuser, die aufgrund ihrer speziellen Nutzung spezifische bauliche und technische Herausforderungen mit sich bringen [1]. So müssen Krankenhäuser nicht nur barrierefrei und für den Krankentransport ausgelegt sein, sondern verfügen zudem über eine umfangreiche technische Gebäudeausstattung – einschließlich Medizin-, Labor- und Notfalltechnik – und müssen darüber hinaus spezifische Anforderungen im Umgang mit Gefahrenstoffen erfüllen. [2].

Diese vielfältigen Anforderungen spiegeln sich in den umfangreichen und komplexen legislativen Rahmenbedingungen des Krankenhausbaus wider. Einen wesentlichen BIM-Anwendungsfall stellt in diesem Kontext die automatisierte Prüfung digitaler Gebäudemodelle auf die Einhaltung gesetzlicher Vorschriften dar. Insbesondere bei Bauwerken, die eine Vielzahl an Anforderungen erfüllen müssen, bietet die automatisierte, regelbasierte Prüfung ein erhebliches Potenzial zur Sicherstellung der Planungsqualität. Solche Prüfverfahren eignen sich auch für den Einsatz im Baugenehmigungsverfahren, da auf diese Weise die Einhaltung baurechtlicher Vorschriften automatisiert durch die zuständigen Behörden überprüft werden kann. Voraussetzung hierfür ist jedoch, dass die baurechtlichen Anforderungen in eine maschinenlesbare Form überführt werden, um sie anschließend automatisiert prüfen zu können. Die Erstellung und Durchführung solcher Prüfregeln erfolgt typischerweise mithilfe spezialisierter Model-Checking-Software. In diesem Beitrag wird dafür die Software Solibri genutzt. Alternative Technologien, wie visuelle Programmierung (z. B. mit Dynamo) oder ontologiegestützte Reasoner [3], werden in diesem Beitrag nicht betrachtet.

Ziel dieser Studie ist die Entwicklung eines Regelsatzes für die automatisierte Qualitätsprüfung im Krankenhausbau mithilfe der Model-Checking-Software Solibri. Der Fokus liegt auf den technischen Brandschutzanforderungen, da insbesondere in diesem Bereich erhebliche Herausforderungen aufgrund komplexer Schnittstellen und hoher Anforderungen an den Patientenschutz bestehen. Grundlage für den Regelsatz bilden die Brandschutzanforderungen aus der Muster-Krankenhausbauverordnung (KhBauVO, 1976). Die zentrale Forschungsfrage lautet: *„Inwieweit lassen sich die technischen Brandschutzanforderungen der KhBauVO in Regelsätze für eine automatisierte Modellprüfung überführen und welche Herausforderungen bestehen bei der Implementierung?“*

Der Beitrag gliedert sich wie folgt: Zunächst werden notwendige Hintergrundinformationen zu den spezifischen Brandschutzanforderungen im deutschen Krankenhausbau sowie zur KhBauVO dargestellt. Anschließend folgt eine Erläuterung des Konzepts der automatisierten Regelprüfung und der verwendeten Software Solibri. Danach wird der Prozess der Regelerstellung und der regelbasierten Prüfung auf Grundlage der KhBauVO detailliert beschrieben und an einem Demonstrator evaluiert. Abschließend werden die Herausforderungen bei der Umsetzung diskutiert, die erzielten Ergebnisse zusammengefasst und ein Ausblick auf weitere Forschungs- und Anwendungsmöglichkeiten gegeben. Die Regelsätze zur Vorab-Prüfung der Klassifikation sowie zur regelbasierten

Überprüfung der Vorschriften mit Solibri sind auf der Plattform GitHub veröffentlicht und können unter https://github.com/SabHartmann/KhBauVO_Brandschutz_Regelsatz eingesehen und verwendet werden.

2 Stand der Technik

Die legislativen Rahmenbedingungen des deutschen Baurechts sind vielfältig und werden nach Bundesland und Art des Gebäudes differenziert. Gemäß Grundgesetz, Art. 30 und Art. 70, obliegt das Bauordnungsrecht dem Landesrecht, wodurch es 16 verschiedene Landesbauordnungen gibt, die sich an der Musterbauordnung (MBO) orientieren. Aufgrund der Art und Nutzung von Krankenhäusern, fallen diese gemäß § 2 Abs. 4 Nr. 10 MBO (2002) in die Kategorie des Sonderbaus und es gelten erhöhte technische Anforderungen, etwa an die Barrierefreiheit oder an den Brandschutz, welche in der KhBauVO definiert sind. Diese ist jedoch nicht rechtsbindend, sondern stellt lediglich eine Empfehlung dar, deren Umsetzung den einzelnen Bundesländern obliegt [4]. Einige Bundesländer haben eigene krankenhausspezifische Sonderbauvorschriften entwickelt. So wurde in Berlin die Krankenhaus-Verordnung (KhsVO, 2012) eingeführt, die Vorgaben zur Flächennutzung und zur Mindestzahl spezieller Räume enthält. Baden-Württemberg verfügt darüber hinaus über ein bauliches Brandschutzkonzept [5], das detaillierte Vorgaben zum Brandschutz im Krankenhausbau enthält – beispielsweise zu den Brandschutzeigenschaften von Materialien, zur Abgrenzung von Brandabschnitten sowie zu Rettungswegen und technischen Einrichtungen. Neben diesen landesspezifischen Sonderbauvorschriften gilt jeweils auch die allgemeine Bauordnung des betreffenden Bundeslandes. Für Bauanträge im Krankenhausbau ist zudem ein projektspezifisches Brandschutzkonzept erforderlich, das individuelle Anpassungen oder Abweichungen von gesetzlichen Vorgaben darlegt [1]. Diese Vielfalt und Individualität der Anforderungen stellt eine große Herausforderung bei der Entwicklung eines einheitlichen Regelsatzes für automatisierte Prüfverfahren dar. Die Transkription der Brandschutzanforderungen aus der KhBauVO in einen maschinenlesbaren Regelsatz dient in dieser Studie daher exemplarisch zur Untersuchung der Machbarkeit automatisierter Prüfverfahren und könnte als Grundlage für erste Modellprüfungen verwendet werden. Eine finale Prüfung sollte jedoch stets projektspezifische und bundeslandspezifische Anforderungen berücksichtigen.

Diese Untersuchung konzentriert sich auf den technischen Brandschutz, der zusammen mit dem organisatorischen Brandschutz dem vorbeugenden Brandschutz zugeordnet wird (s. Abbildung 1). Zum technischen Brandschutz zählen unter anderem die Prüfung von Nutzungseinheiten, die Klassifizierung von Baustoffen hinsichtlich ihrer Brennbarkeit sowie die Bewertung der Feuerwiderstandsfähigkeit von Bauteilen. In der KhBauVO werden diese Anforderungen speziell in den §§ 4, 7 bis 18 sowie 22, 24 bis 26 und 31 adressiert.

Für die Durchführung der regelbasierten Prüfung müssen die relevanten Informationen zunächst in das Gebäudemodell implementiert und in ein geeignetes Format, wie Industry Foundation Classes (IFC), exportiert werden. Verschiedene Model-Checking-Softwarelösungen ermöglichen die automatisierte Prüfung solcher Modelle. *Hjelseth* [7] gibt einen Überblick über solche Lösungen und stellt fest, dass Solibri im Vergleich zu anderen Model Checkern über umfangreiche Prüfmöglichkeiten (Attribut- und Kollisionsprüfung sowie das Compliance Checking) und eine IFC-Schnittstelle verfügt und z. B. in

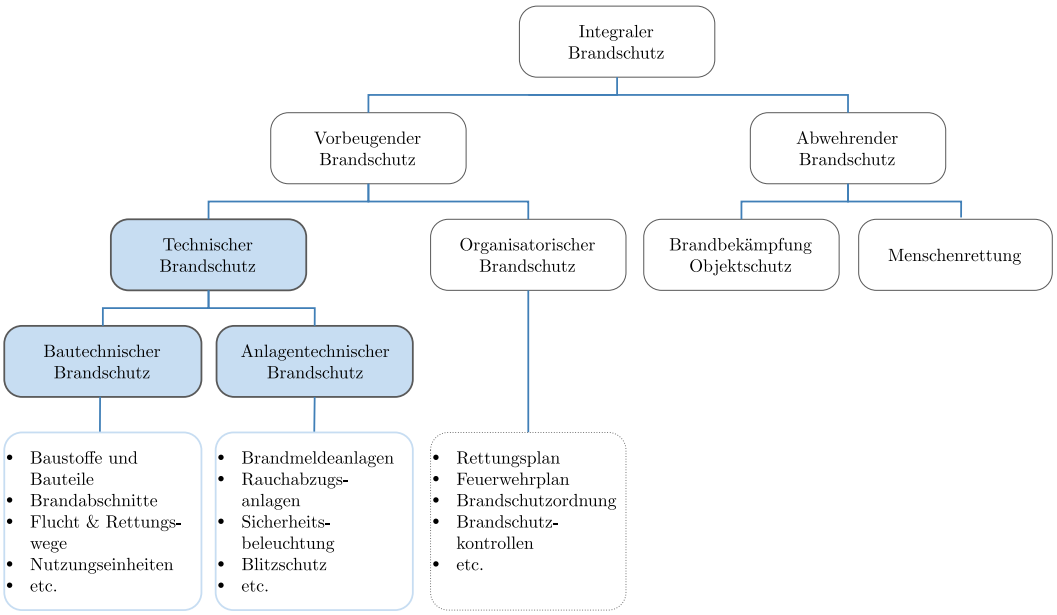


Abbildung 1: Struktur des integralen Brandschutzes nach Weller [6]

Finnland als Prüfwerkzeug im Baugenehmigungsprozess eingesetzt wird. Aufgrund der Verfügbarkeit und Funktionsvielfalt wird in dieser Studie die Software Solibri verwendet.

Die regelbasierte Prüfung erfolgt nach [8] in folgenden Schritten: Ein BIM-Modell wird zunächst in einer Autorensoftware erstellt und anschließend als IFC-Datei exportiert. Gegebenenfalls erfolgt ein Zwischenschritt über eine Common Data Environment (CDE), um das Modell innerhalb des Projektteams zu teilen. Anschließend erfolgt die Prüfung in der Model-Checking-Software – in diesem Fall Solibri – welche die Einhaltung der vordefinierten Regeln überprüft. Die Prüfergebnisse, die aufgedeckte Fehler oder Unstimmigkeiten enthalten, können im BIM Collaboration Format (BCF) exportiert und in der Autorensoftware angepasst werden. Der beschriebene Arbeitsablauf ist in Abbildung 2 dargestellt.

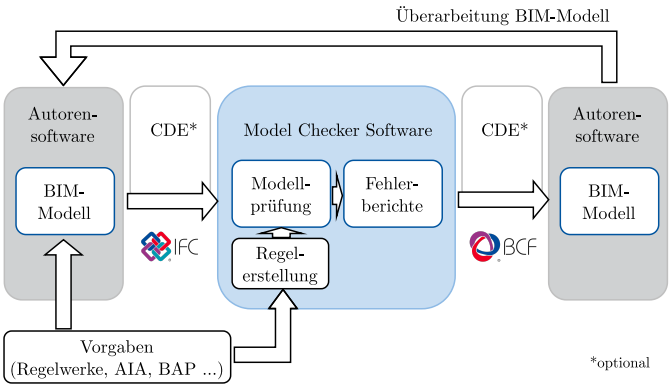


Abbildung 2: Regelbasierter Prüfprozess in Solibri

3 Durchführung der Modellprüfung

3.1 Modellvorbereitung

Eine regelbasierte Modellprüfung setzt voraus, dass das Gebäudemodell semantisch korrekt attribuiert ist. Dazu müssen alle relevanten Informationen der Gebäudeelemente in Form von Properties bzw. Property Sets definiert und exportiert werden. In dieser Studie wird die Autorensoftware Revit verwendet, um ein vereinfachtes Gebäudemodell zu erstellen, das – obwohl nicht explizit als Krankenhaus modelliert – sämtliche relevanten Brandschutzinformationen enthält. Diese Informationen werden gemäß den Handlungsempfehlungen des Vereins zur Förderung von Ingenieurmethoden im Brandschutz (VIB) implementiert [9], welcher die IFC-Attribute hinsichtlich deutscher Brandschutzanforderungen erweitert. Beispielsweise empfiehlt der VIB etwa, die Baustoffklasse nicht als Boolean-Wert *„Combustible“* zu definieren, sondern im Textformat mit detaillierten Klassen (A1 bis B3) zu hinterlegen.

Zusätzlich werden weitere Attribute definiert, die für die Prüfung technischer Brandschutzanforderungen gemäß der KhBauVO erforderlich sind (s. Tabelle 1). Diese Attribute werden in Revit als gemeinsam genutzte Parameter angelegt, was eine konsistente Attribuierung und damit die nachfolgende regelbasierte Prüfung erleichtert. Weiterhin wird auf eine korrekte IFC-Klassifizierungen geachtet, sodass beispielsweise Dämmschichten nicht standardmäßig als Wände *IfcWall*, sondern spezifisch als *IfcCovering – INSULATION* klassifiziert werden.

Tabelle 1: Ergänzende Attribute für den Brandschutz (* dichtschließend (dt) / dicht- und selbstschließend (dts) / vollwandig, dicht- und selbstschließend (vdts) / rauchdicht (rd))

Parameter	Attributwerte	Bauteilkategorie	Beschreibung
FSR_clamping_device	ja/nein	Sonderausstattung	Haltevorrichtung in Aufzügen
FSR_horizontal_pressure	Zahl [kN/m]	Geländer/Wand	Gibt die horizontale Druckwiderstandsfähigkeit an
FSR_tight-/self-closing_door	dt/dts/vdts/rd*	Türen	Rauchdichte und selbstschließende Türen
FSR_door_drive	ja/nein	Türen	Automatischer Türantrieb
FSR_nonskid_material	ja/nein	Rampen/Räume/ Sonderausstattung	Rutschfester Boden

3.2 Implementierung der KhBauVO in Solibri

Nach der Modellvorbereitung in Revit erfolgt die Regelerstellung in Solibri. Dazu wird zunächst eine spezifische Rolle *„KhBauVO“* erstellt, welcher der Regelsatz *Muster-Krankenhausbauverordnung.cset* sowie die erforderlichen Klassifikationen zugeordnet werden. Die Regelerstellung in Solibri basiert auf Standardregelvorlagen, welche spezifisch angepasst werden können. Jede Regel umfasst typischerweise zunächst eine Auswahl der zu überprüfenden Komponenten, die anschließend anhand numerischer Werte oder spezifischer Attribute validiert werden. Die Auswahl erfolgt entweder über vordefinierte Bauteilkategorien, IFC-Klassifikationen oder benutzerdefinierte Klassifikationen. Letztere ermöglichen eine präzisere Gruppierung von Elementen, etwa die Unterteilung von Wänden in Außen-

und Innenwände oder tragende und nicht-tragende Bauteile anhand der IFC-Attribute *IsExternal* und *LoadBearing*.

Im nächsten Schritt erfolgt die Regelerstellung mithilfe des *RuleSet Managers* in Solibri. Jede relevante Vorschrift der KhBauVO durchläuft zunächst eine Vorabprüfung hinsichtlich korrekter Attribuierung der verwendeten Komponenten, wofür entsprechende Klassifikationsregeln zu erstellen sind. Durch diese strukturierte Auswertung lassen sich Fehlerursachen leichter identifizieren. Erst dann erfolgt die eigentliche Prüfung auf Einhaltung der Anforderungen. Die Prüfung eines einzelnen Absatzes setzt sich dabei oft aus mehreren, hierarchisch aufgebauten Teilregeln zusammen. Ein Beispiel zeigt § 7 Abs. 1 der KhBauVO, der fordert, dass tragende und aussteifende Bauteile in Gebäuden mit mehreren Geschossen feuerbeständig sein müssen. Diese Anforderung wird zunächst anhand der Gebäudehöhe überprüft, sodass erst im nächsten Schritt die Feuerbeständigkeit der betroffenen Bauteile geprüft wird (s. Abbildung 3). An die Baustoffklasse werden nach § 7 Abs. 1 KhBauVO keine klaren Anforderungen gestellt, jedoch dürfen "... Bedenken wegen des Brandschutzes nicht bestehen". Diese Formulierung lässt Interpretationsspielraum zu und kann daher nicht allgemeingültig geprüft bzw. als maschinenlesbare Regel abgebildet werden.

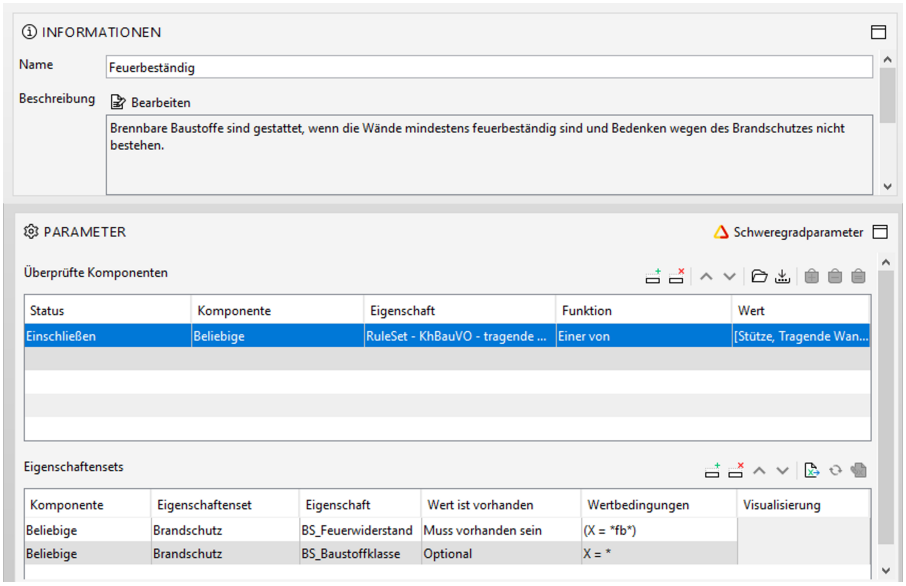


Abbildung 3: Teilprüfung der erforderlichen Feuerbeständigkeit (fb) bestimmter Bauteile

Die erstellten Regeln werden anschließend auf das vorbereitete Gebäudemodell angewendet, um ihre Praxistauglichkeit zu evaluieren. Bei der Prüfung von § 14 Abs. 3 bis 6, der bauliche Anforderungen an Treppen behandelt, ergeben sich beispielsweise verschiedene Fehlermeldungen – etwa zu geringe Treppenbreiten oder variierende Stufenhöhen. Ein weiteres Beispiel ist die Überprüfung der Rauchabzüge gemäß § 15 Abs. 3, die auf eine falsche Positionierung im Modell hinweist. Die Prüfung der Brandabschnitte nach § 10 Abs. 1 erweist sich hingegen als nur eingeschränkt praktikabel. Zwar können Räume über eine entsprechende Attribuierung in der Autoren-Software zu Brandabschnitten gruppiert werden, jedoch stehen in Solibri für diese Gruppen keine weiterführenden Regelprüfungen zur Verfügung. Eine Überprüfung ist lediglich visuell möglich, ergänzt um die Ausgabe der Bruttofläche.

Zudem werden geschossübergreifende Brandabschnitte auf diese Weise nicht als zusammenhängend erkannt, sondern in mehrere Einzelabschnitte aufgeteilt. Eine alternative Herangehensweise besteht in der Zoneneinteilung in Revit, die in Solibri einer Regelprüfung unterzogen werden kann. Dabei zeigt sich jedoch, dass Solibri Räume innerhalb einer Zone isoliert bewertet. Infolgedessen werden Räume, die keine direkte Verbindung zu anderen Brandabschnitten oder zu einem Treppenraum aufweisen, fälschlicherweise als fehlerhaft markiert. Da Solibri Zonen nicht ganzheitlich analysiert, sondern ausschließlich auf Raumebene prüft, lassen sich zusammenhängende Brandabschnitte nicht zuverlässig bewerten. Aus diesem Grund wird empfohlen, die Zonenzuordnung direkt in Solibri vorzunehmen und die Überprüfung manuell durchzuführen.

4 Diskussion

Die manuelle semantische Anreicherung des Modells, die Klassifikation der Parameter und die Erstellung der Prüfregeln sind zeitaufwendige Prozesse. Die Anwendung des erstellten Regelsatzes auf ein beispielhaftes BIM-Modell zeigt jedoch, dass typische Planungsfehler im Brandschutz automatisiert erkannt werden können, wie z. B. widersprüchliche Treppenmaße. Gleichzeitig offenbaren sich bei der praktischen Umsetzung verschiedene Herausforderungen. Damit erschwert die heterogene Gesetzgebung die Implementierung automatisierter Regelprüfungen erheblich, da individuelle Vorschriften der Bundesländer sowie projektspezifische Gutachten im Krankenhausbau zu berücksichtigen sind. Dadurch sind spezifische Anpassungen der Regelsätze erforderlich, die zeitintensiv sind und vertiefte Kenntnisse der verwendeten Software bzw. Regellogik voraussetzen. Auch die eingesetzten Softwarelösungen bringen Einschränkungen mit sich. Eine konsistente Attribuierung in Revit ist die Voraussetzung für eine anschließende Prüfung, welche jedoch teilweise manuell erfolgen muss und fehleranfällig ist. Zudem gestaltet sich in Solibri die Anwendung vordefinierter Regelsätze kompliziert, da Nutzer diese nur eingeschränkt modifizieren können. Darüber hinaus enthalten viele Vorgaben unspezifische Formulierungen, wie „in Türnähe“ oder „günstig gelegen“, die sich aufgrund von qualitativen Aussagen nicht objektiv prüfen und in eine maschinenlesbare Prüflogik überführen lassen. Qualitative Anforderungen enthalten nicht messbare Vorgaben und müssen durch Experten nach fachlichem Ermessen beurteilt werden. Um den nächsten Schritt Richtung praktischer Anwendung zu ermöglichen, gilt es zunächst, die Methodik und den Regelsatz an einem realen Krankenhaus zu validieren und die Regelbasis um weitere Fachbereiche zu vervollständigen.

5 Fazit und Ausblick

Der entwickelte Regelsatz auf Basis der KhBauVO bietet eine geeignete Grundlage für die automatisierte Prüfung brandschutztechnischer Anforderungen im BIM-Prozess. Für eine umfassendere Anwendung und höhere Effizienz ist der Ansatz jedoch in verschiedenen Bereichen weiter zu optimieren. So ist eine vollständige Integration legislativer Anforderungen weiterer Fachbereiche erforderlich, um den Regelsatz umfassend nutzbar zu machen. Darüber hinaus könnten Verbesserungen in den Softwarelösungen – etwa durch automatische Parameterbefüllung z. B. mittels Dynamo-Skripten in Revit oder durch standardisierte IFC-Parameter – sowohl den Vorbereitungsaufwand als auch potenzielle Fehlerquellen deutlich reduzieren. Eine stärkere Anbindung von Materialbibliotheken würde die automatische Verknüpfung von Bauteilen mit brandschutzrelevanten Eigenschaften ermöglichen und damit

eine automatische Zuweisung der Baustoffklasse im Modell unterstützen. Einheitliche gesetzliche Standards auf Bundesebene wären ein wesentlicher Schritt, um den Aufwand der Regelset-Anpassung zu minimieren. Die Einführung eines modellbasierten Genehmigungs- und Prüfprozesses – analog zum Vorgehen in Finnland – könnte schließlich dazu beitragen, die Effizienz in der Planung zu steigern, die Planungssicherheit zu erhöhen und dem Fachkräftemangel im Bauwesen wirksam entgegenzuwirken. Eine alternative Herangehensweise wäre der Einsatz Künstlicher Intelligenz, um maschinenlesbare Regeln automatisiert auf Grundlage von Gesetzestexten zu generieren. Die Ergebnisse wären jedoch auch weiterhin fachlich zu überprüfen, sodass die Haftung weiterhin beim menschlichen Entscheider verbleibt und Künstliche Intelligenz lediglich ein unterstützendes Werkzeug bei der Regelerstellung bietet Insgesamt bietet die automatisierte Modellprüfung somit großes Potenzial zur Förderung der Digitalisierung und stellt ein zukunftsweisendes Instrument für Planer und Unternehmen dar.

Literatur

[1] S. Hartmann, M. Abodewan und K. Klemt-Albert, »Intransparente Vielfalt im Krankenhausbau«, 33. BBB-Assistent:innentreffen, Aachen (Germany), 15 Jul 2024 - 17 Jul 2024, RWTH publications, 2024, S. 95–109. DOI: 10.18154/RWTH-2024-09723

[2] C. Lange, *Vorbeugender Brandschutz in der Gebäudeplanung: Vom Entwurf zum Brandschutzkonzept*. Wiesbaden: Springer Vieweg, 2021. DOI: 10.1007/978-3-658-34932-5

[3] S. Li, Z. Jiang und Z. Xu, »BIM-Based Model Checking: A Scientometric Analysis and Critical Review«, *Applied Sciences*, Jg. 15, Nr. 1, 2025. DOI: 10.3390/app15010049

[4] A. Merschbacher, »Flucht- und Rettungswege - Anforderungen behinderter Menschen an die Bewältigung von Notfällen«, 2021. DOI: 10.1007/978-3-658-32845-0

[5] Wirtschaftsministerium Baden-Württemberg, *Hinweise des Wirtschaftsministeriums Baden-Württemberg über den baulichen Brandschutz in Krankenhäusern und baulichen Anlagen entsprechender Zweckbestimmung*, https://mlw.baden-wuerttemberg.de/de/bauen-wohnen/baurecht/erlasse-und-vorschriften.

[6] B. Weller und S. Heilmann, »Brandschutz«, in *Ulrich Vismann (Hg.) - 2021 - Wendehorst Bautechnische Zahlentafeln*, 2021, S. 1123–1156. DOI: 10.1007/978-3-658-32218-2

[7] E. Hjelseth, »BIM-based model checking (BMC)«, *Building Information Modeling: Applications and Practices*, S. 33–61, Mai 2015. DOI: 10.1061/9780784413982.ch02

[8] A. Speth, *Solibri Schulung - Schulungsunterlagen*, Online, 2023.

[9] Verein zur Fördererung von Ingenieurmethoden im Brandschutz, *Handlungsempfehlung Digitalisierung und Building Information Modeling für die Brandschutzplanung*, https://www.vib-brandschutz.de/fileadmin/media/dokumente/Nov-2023-Handlungsempfehlung-Digitalisierung-und-BIM-fuer-die-Brandschutzplanung-V1.1.pdf.

Synthetic Image Data from 3D Software Environments and Game Engines for Training Machine Learning Models in Damage Detection

Mahmoud Moubarak and Baris Özcan
Geodetic Institute and Chair for Computing in Civil Engineering & Geo Information Systems,
RWTH Aachen University, Mies-van-der-Rohe-Str. 1, 52074 Aachen, Germany
E-mail(s): mahmoud.moubarak@rwth-aachen.de, oezcan@gia.rwth-aachen.de

Abstract: The development of supervised machine learning (ML) models for the detection of concrete damage requires large amounts of labeled image data. Such data usually originates from real images, but collecting them is time-consuming, expensive, error-prone and logistically difficult. The use of synthetically generated images may offer a viable alternative to supplement or even partially replace real data. In this paper, we investigate the potential use of synthetic images to improve ML models for crack detection in concrete surfaces. The images are generated using 3D software and game engines such as Unreal Engine, Blender and different video games. Unreal Engine delivered good results due to static textures and high visual quality, but is limited in variation. Blender performed slightly weaker, but enabled unlimited and diverse image generation through procedural textures. Capturing screenshots from video games initially showed potential, but proved to be less effective overall. In general, our results show that 3D software environments are suitable for generating synthetic training data for this domain.

Keywords: Crack Detection, Deep Learning, Synthetic Data Generation, Unreal Engine, Blender

DOI: 10.18154/RWTH-CONV-254919. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany,
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

Detecting structural damage such as concrete cracks is time-consuming, costly, error-prone, and often logistically difficult. Image-based machine learning (ML) approaches, which require large amounts of high-quality image training data, are increasingly being used for automation. However, acquiring diverse training data is challenging, as damage occurs under varying conditions and its documentation is often limited and time-consuming. Existing datasets often only cover a small proportion of relevant variants, which makes it difficult to develop robust models.

A promising approach is the use of synthetically generated image data. With 3D simulations or game engines, realistic damage scenarios can be generated tailored to the purpose, for example by varying crack shapes, materials and environmental conditions. In this way, synthetic data can usefully supplement or partially replace real datasets.

In this paper, we examine the suitability of synthetically generated image data from 3D environments and game engines for training ML models for concrete crack detection. For this purpose, various damage scenarios have been modeled, visually rendered and specifically varied in terms of crack shapes, material properties, lighting conditions and backgrounds. The objective is to utilize the generated data as a supplement to real-world images and to evaluate their effect on model performance and robustness in order to demonstrate the contribution of synthetic data to improving model reliability.

The remainder of this paper is structured as follows: Section 2 presents a brief overview of related work on damage detection and synthetic data for ML models. In Section 3, we describe our methods to generate synthetic training data. The experimental results are presented in Section 4 and discussed in Section 5. Finally, we conclude in Section 6 with an additional outlook.

2 Related Scientific Work

The following section provides an overview of relevant work on automated damage detection and the use of synthetic image data in ML.

2.1 Automated Damage Detection

In recent years, substantial progress has been made in the automated detection of damage. Damage to concrete structures is now often analyzed using image processing techniques and deep neural networks [1], particularly Convolutional Neural Networks (CNNs) trained on large datasets.

Several studies have investigated the use of ML-based approaches for infrastructure damage detection. Among these, pre-trained CNNs such as ResNet-50 have been highlighted for their advantages [2]. Transfer learning methods were also explored, with various pre-trained models assessed for their suitability in this domain. EfficientNetB0 achieved the best results due to its balanced characteristics in model size, speed, and accuracy [3]. The strengths of the EfficientNet series have also been confirmed in subsequent research [4].

2.2 Synthetically Generated Image Data

The use of synthetically generated image data is becoming increasingly important in ML due to the limited availability of real-world training data. For instance, image augmentations such as rotation, displacement and scaling of real images can improve model performance [5], even if suitable transformations usually have to be determined empirically.

Current research also employs game and modeling engines to generate synthetic training data. For example, Unity has been used to simulate natural ice environments, where 65% of the data were generated synthetically and still yielded an Intersection over Union of 83% despite domain shifts [6]. Unreal Engine 4 has been applied to visualize crack and rust structures on infrastructure surfaces, with a ResUNet trained on combined real and synthetic data achieving F1 scores of 89.2% for cracks and 84.1% for rust [7]. Blender has been utilized to create synthetic steel damage via procedural textures, where a U-Net trained on synthetic images outperformed its real-data counterpart (dice score 0.63 vs. 0.56) [8]. Comparable segmentation results have also been reported for synthetic vehicle damage images [9].

3 Methodology

In this section, three different methods for the synthetic generation of image data of concrete cracks are described, which are based on Unreal Engine, Blender and video games.

3.1 Unreal Engine

For our first approach, we used Unreal Engine 5.4.1 [10] to generate synthetic image data as it supports the representation of realistic surfaces. In this setup, static crack textures were projected onto 3D models using UV coordinates, which determine the placement of the textures on a surface. The U and V values are normalized to a range from 0 to 1, with values outside this range leading to a repetition of the texture. The variation of the textures was increased through specific transformations such as shifting, scaling and rotation. Parameters allowed for automated adjustment of the appearance and contributed to the diversity of the generated surfaces. This was crucial to avoid overfitting in model training.

Three texture types were used to represent concrete surfaces (see Figure 1):

- Diffuse textures for base color (RGB)
- Normal textures for depth effect from altered surface normals
- ORDp textures for occlusion, roughness, metalness (only roughness used)

Most of the textures were imported via Quixel Bridge, Epic Games' library for numerous textures, including realistic surfaces and materials, and adjusted in the material editor.

Figure 1: Left: Three-dimensional surface perception of a normal texture. Center: Materials. Right: Three types of textures: Diffuse, Normal, and ORDp.

Blueprint scripts were used to create 10×10 unit concrete walls for data generation. Random parameters for scaling, rotation and color were used to create different patterns based on the same texture. The value ranges were determined experimentally. The lighting conditions were varied manually to simulate different environmental conditions. Finally, high-resolution images (2048×2048 pixels) were taken with an orthogonally aligned camera. In total, 700 images of intact and 700 images of cracked concrete surfaces have been collected (see Figure 2), including realistic details such as dirt or soiling.

3.2 Blender

The second approach to generating synthetic image data uses the 3D modeling software Blender [11]. In contrast to the Unreal Engine, the dataset here is based entirely on procedural textures, which do not consist of static images but are calculated dynamically using mathematical functions. This allows an unlimited number of variants to be generated, in particular individual crack structures, which creates a robust and diverse data basis for training AI models.

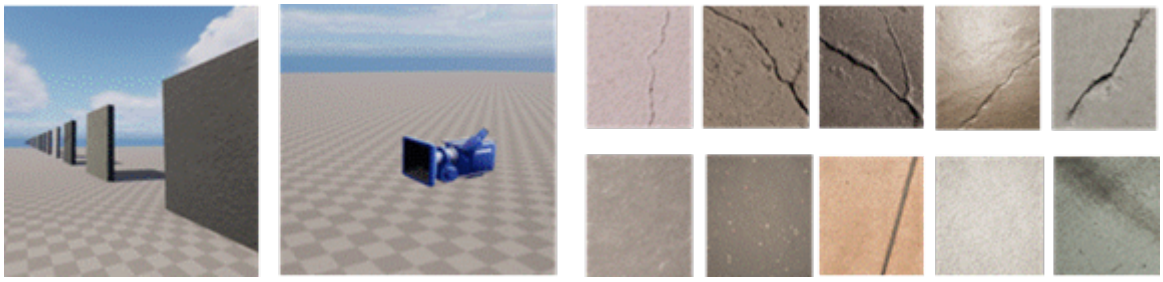


Figure 2: Rendering process and concrete surface results from Unreal Engine

A common method is the Voronoi texture, which divides the surface into cells assigned to the nearest point (Voronoi sites), with distances calculated using the Euclidean metric. On its own, this produces only cell-like structures; however, when combined with noise textures, the resulting patterns begin to resemble cracks more closely. With the aid of the Distance to Edge function integrated in Blender, cell boundaries can be specifically emphasized, creating mosaic-like, realistic-looking fracture edges (see Figure 3).

To increase structural complexity, noise textures are used (see Figure 3). While these generate small-scale irregularities, they still lack coherent crack patterns. This limitation can be addressed through Spectral Synthesis, where multiple noise layers are superimposed. When Perlin noise is used as the base function, this process is referred to as fractional Brownian motion (fBm) (1) [12], producing structures that combine both fine and large-scale details.

The fBm function is described as follows:

$$f_s(x) = \sum_i \omega_i \cdot f(s_i \cdot x) \tag{1}$$

where ω_i are the weights, s_i are the scalings and $f(s_i \cdot x)$ is the noise function at the respective frequency. The result is a weighted sum of all layers, which forms the final texture.

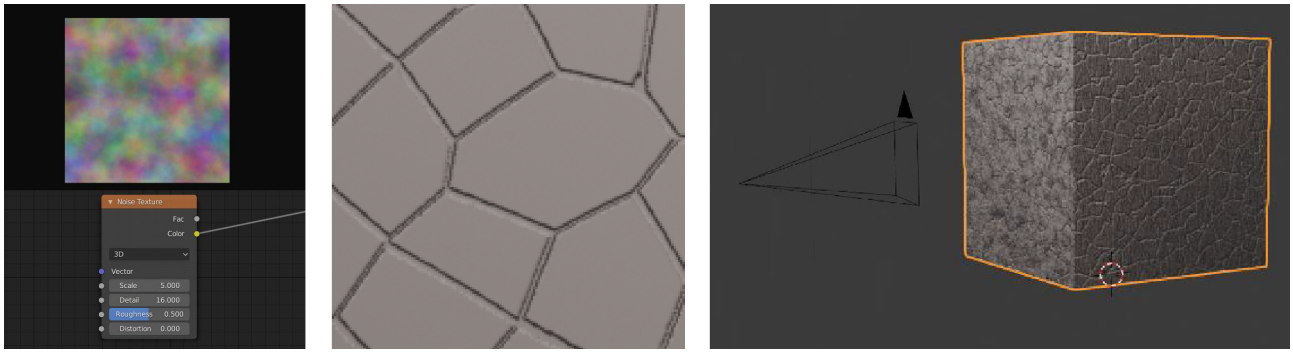


Figure 3: Noise Texture, Voronoi Texture, and rendering process

Blender supports up to 15 octaves with customizable parameters, allowing precise control over the resulting patterns. In addition, the bump mapping process is used to change surface normals locally using the procedural texture. The geometry remains unchanged, but a realistic depth effect is created by light reflections on edges and cracks.

The crack structures are randomly varied using a Python script and the images are rendered in real time using the Eevee engine. This creates a great visual variety with a realistic representation of damaged concrete surfaces (see Figure 4).

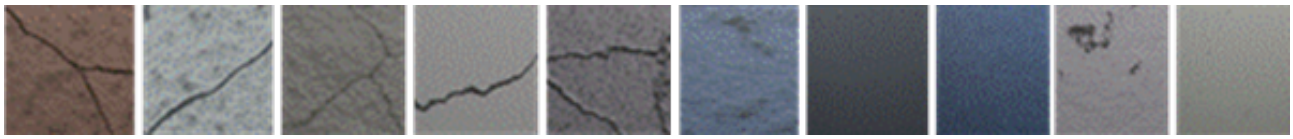


Figure 4: Concrete surface results from Blender

3.3 Video games

The third approach involved taking screenshots from various video games. As the games were based on different game engines, this resulted in a collection of images of varying quality. Most of the games come from free and publicly accessible sources. Accordingly, some older engines were involved, so that the image quality does not meet current graphic standards. Some examples of images are shown in Figure 5.

Figure 5: Concrete surface results from capturing screenshots from video games

4 Experimental Results

Five datasets with 1400 training images each were used for the experiments:

- the real-world dataset “Concrete Crack Images for Classification” (CCIC) [13, 14],
- the real-world dataset “SDNET2018” [15, 16],
- a synthetic dataset based on Unreal Engine,
- a synthetic dataset based on Blender and
- a synthetic dataset from video games.

For each dataset, one model was trained with pre-trained ImageNet weights and one with random initialization, resulting in 10 models. In addition, seven further randomly initialized models were trained on mixed data. Six of these combined 60% real-world and 40% synthetic images, while another consisted solely of an equal mixture of Unreal Engine and Blender images. This resulted in a total of 17 models, each of which was evaluated using 500 test images from the two real-world datasets (CCIC and SDNET2018). The results are shown in Figure 6.

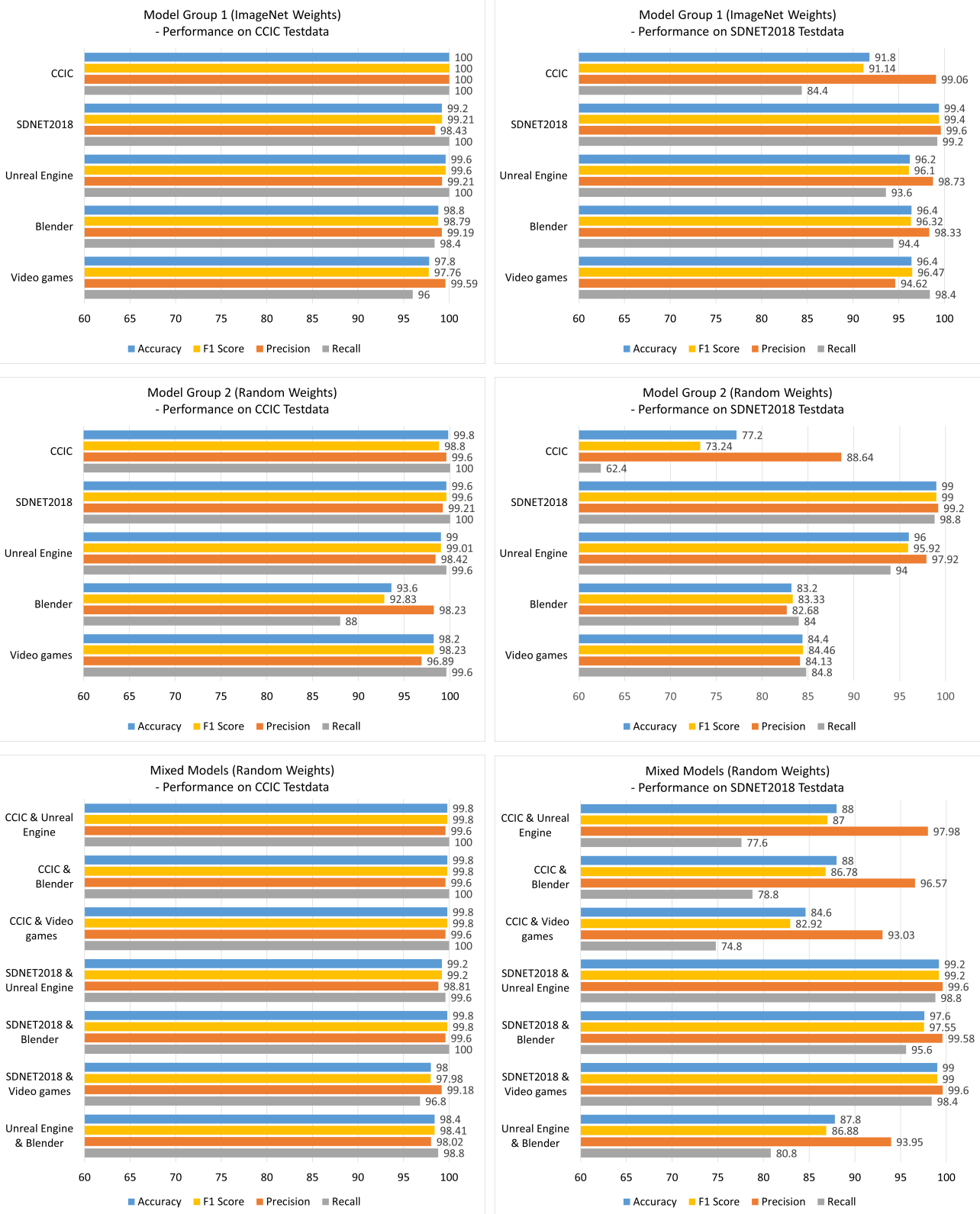


Figure 6: Comparison of the different models on the CCIC and SDNET2018 test data

5 Discussion

A significant increase in performance through synthetic data was especially evident in the SDNET2018 dataset. The model trained exclusively with CCIC data achieved only an F1 score of 73.24%. Adding synthetic images increased this value to 87.00% (Unreal Engine), 86.78% (Blender) and 82.92% (video games). On the CCIC dataset, by contrast, hardly any improvement was visible, as the model already reached very high values with real-world training data. With ImageNet pre-initialization, F1 scores were above 96% in almost all cases, regardless of the dataset. Here, the influence of the training data was much lower than with random initialization.

The type of synthetic source affected both performance and time effort. Unreal Engine provided high-quality images with clearly recognizable crack structures, but was limited in variance and data volume due to static textures. Blender used procedural textures, allowing greater variability in surface appearance, even if crack representation was less clear. Images from video games were easily accessible, but less varied. While video games were most time-consuming, Blender proved most efficient, producing unique crack structures without repetition and with lower effort. Unreal Engine was more demanding due to static textures that had to be downloaded and searched.

6 Conclusion

In this paper, we demonstrated that synthetically generated image data can effectively complement the training of crack detection models. Especially in combination with real training data, they led to improved model performance. The results illustrate that images from Unreal Engine and Blender offer a high potential, as they contribute to the robustness of the training either through graphical quality or procedural variance. Synthetic data can therefore be a useful extension of real datasets, especially when their quantity is limited.

Future work could focus on the automatic generation of realistic damage patterns by integrating additional plugins or software such as Substance 3D Designer in Unreal Engine. In addition, the use of synthetic data for more complex types of damage and for tasks such as object recognition and segmentation could be investigated. Targeted post-processing of images using filters also offers further potential for optimization.

References

[1] B. Özcan, D. Crampen, Z. Kratzer, and J. Blankenbach, “Holistic image-based analysis of damage on concrete surfaces—a multifaceted approach based on supervised machine learning”, *Structural Concrete*, vol. 25, no. 5, pp. 3121–3144, 2024. DOI: 10.1002/suco.202300626 [Online]. Available: https://doi.org/10.1002/suco.202300626

[2] A. Kumar, H. Martin, and L. Leon. “Concrete damage identification for structural health monitoring using computer vision”. [Online]. Available: https://doi.org/10.5281/ZENODO.8066307

[3] C. Su and W. Wang, “Concrete cracks detection using convolutional neuronal network based on transfer learning”, *Mathematical Problems in Engineering*, 2020. DOI: 10.1155/2020/7240129 [Online]. Available: https://doi.org/10.1155/2020/7240129

[4] S. S. Zadeh, S. Aalipour Birgani, M. Khorshidi, and F. Kooban. “Concrete surface crack detection with convolutional-based deep learning models”. [Online]. Available: https://doi.org/10.5281/ZENODO.10061654

[5] A. Mumuni and F. Mumuni, “Data augmentation: A comprehensive survey of modern approaches”, *Array*, vol. 16, p. 100 258, 2022. DOI: 10.1016/j.array.2022.100258 [Online]. Available: https://doi.org/10.1016/j.array.2022.100258

[6] S. Taheri and V. Vardanyan. “The power of synthetic data: Using game engines to improve ai”. [Online]. Available: https://medium.com/cdrin-cctt/the-power-of-synthetic-data-using-game-enginesto-improve-ai-67e872bab997

[7] L. A. Zherdeva, E. Y. Minaev, and N. A. Firsov, “Building surface damage recognition based on synthetic data”, in *2023 IX International Conference on Information Technology and Nanotechnology (ITNT)*, IEEE, 2023, pp. 1–6. DOI: 10.1109/ITNT57377.2023.10139288 [Online]. Available: https://doi.org/10.1109/ITNT57377.2023.10139288

[8] A. Boikov, V. Payor, R. Savelev, and A. Kolesnikov, “Synthetic data generation for steel defect detection and classification using deep learning”, *Symmetry*, vol. 13, no. 7, p. 1176, 2021. DOI: 10.3390/sym13071176 [Online]. Available: https://doi.org/10.3390/sym13071176

[9] J. Parslov, E. Riise, and D. P. Papadopoulos, “Crashcar101: Procedural generation for damage assessment”, *arXiv preprint arXiv:2311.06536*, 2023. DOI: 10.48550/ARXIV.2311.06536 [Online]. Available: https://doi.org/10.48550/ARXIV.2311.06536

[10] Epic Games. “Unreal engine documentation”. [Online]. Available: https://docs.unrealengine.com

[11] Blender Foundation. “Blender documentation”. [Online]. Available: https://docs.blender.org

[12] K. Perlin, “An image synthesizer”, *Computer Graphics*, vol. 19, no. 3, pp. 287–296, 1985. DOI: 10.1145/325165.325247 [Online]. Available: https://doi.org/10.1145/325165.325247

[13] Ç. F. Özgenel and A. Gönenç Sorguç, “Performance comparison of pretrained convolutional neural networks on crack detection in buildings”, in *Proc. of ISARC 2018*, Berlin, 2018.

[14] Ç. F. Özgenel, *Concrete crack images for classification*, version V2, Mendeley Data, 2019. DOI: 10.17632/5y9wdsg2zt.2 [Online]. Available: https://doi.org/10.17632/5y9wdsg2zt.2

[15] S. Dorafshan, R. J. Thomas, and M. Maguire, “SDNET2018: An annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks”, *Data in Brief*, vol. 21, pp. 1664–1668, 2018. DOI: 10.1016/j.dib.2018.11.015 [Online]. Available: https://doi.org/10.1016/j.dib.2018.11.015

[16] M. Maguire, S. Dorafshan, and R. J. Thomas, *SDNET2018: A concrete crack image dataset for machine learning applications*, 2018. DOI: 10.15142/T3TD19 [Online]. Available: https://digitalcommons.usu.edu/all_datasets/48/

Automatisierte Fehlererkennung im Schmelzschichtverfahren mittels künstlicher Intelligenz

Thamer Al-Zuriqat, Mahmoud Noufal, Adel Alatassi und Jasper Vollmert
Institut für Digitales und Autonomes Bauen, Technische Universität Hamburg, Blohmstraße 15, 21079 Hamburg, Deutschland
E-Mails: thamer.al-zuriqat@tuhh.de

Abstract: Das Schmelzschichtverfahren (engl. Fused Filament Fabrication, FFF) gehört zu den am häufigsten verwendeten additiven Fertigungstechniken und überzeugt durch seine Vielseitigkeit, Kosteneffizienz und Flexibilität. FFF-Maschinen, insbesondere 3D-Drucker, schmelzen ein Filament aus thermoplastischem Material durch Erhitzen und extrudieren das geschmolzene Filament schichtweise durch eine Düse. Allerdings sind FFF-Maschinen anfällig für Laufzeitfehler, die zu Zeit- und Materialverlusten führen können und eine kontinuierliche Überwachung durch Menschen erfordern. In dieser Arbeit wird ein Ansatz zur automatisierten Fehlererkennung (engl. Automated Defect Detection, ADD) für FFF-Prozesse vorgestellt, der faltende neuronale Netze und Bildverarbeitungstechniken kombiniert. Zur Validierung des ADD-Ansatzes wurden gezielt 3D-Drucke erstellt, die typische FFF-Defekte nachbilden, sowie fehlerfreie Druckprodukte verwendet. Die Ergebnisse zeigen, dass der vorgeschlagene Ansatz in der Lage ist, FFF-Defekte automatisch zu erkennen, wodurch Zeit- und Materialverluste sowie die Notwendigkeit menschlicher Überwachung reduziert werden.

Keywords: Additive Fertigung, Fehlererkennung, Fused Filament Fabrication, Künstliche Intelligenz, Computer Vision

DOI: 10.18154/RWTH-CONV-254871. Erschienen in Tagungsband 36. Forum Bauinformatik 2025, Aachen, Deutschland.
© 2025 Das Copyright für diesen Beitrag liegt bei den Autoren. Diese Veröffentlichung, ausgenommen Zitate und anderweitig gekennzeichnete Teile, ist unter der Creative-Commons-Lizenz CC BY 4.0 lizenziert.

1 Einleitung

Das Schmelzschichtverfahren (engl. Fused Filament Fabrication, FFF), auch als 3D-Druck bekannt, gehört zu den am weitesten verbreiteten additiven Fertigungsprozessen. FFF-Maschinen (nachfolgend als „3D-Drucker“ bezeichnet) verwenden Schrittmotoren und Extruder, um polymerbasierte Materialien („Filament“) schichtweise in vordefinierte Formen zu extrudieren und so Produkte („gedruckte Produkte“) zu erzeugen. 3D-Drucker sind anfällig für Laufzeitfehler, die sich in Form von Defekten in den gedruckten Produkten manifestieren können (nachfolgend als „FFF-

Defekte“ bezeichnet). FFF-Defekte umfassen Unregelmäßigkeiten, Imperfektionen und strukturelle Fehlstellen, die die Qualität der gedruckten Produkte beeinträchtigen können. Solche Defekte führen oft zu Zeit- und Materialverlusten, da der FFF-Prozess im Falle eines Fehlers häufig von vorne begonnen werden muss. Daher ist eine kontinuierliche menschliche Überwachung des FFF-Prozesses unerlässlich. FFF-Defekte können sowohl auf den äußeren Oberflächen der gedruckten Produkte auftreten als auch im Inneren entstehen und unbemerkt bleiben.

Ähnlich wie in anderen Bereichen, wie der Schadensdetektion in Strukturen [1], Leichtbau-Brücken [2] oder der Diagnose von Sensorfehlern [3], wurden auch für die automatisierte Fehlererkennung in FFF-Prozessen verschiedene Ansätze erforscht. In [4] wird ein auf Bildverarbeitung basierender Offline-Ansatz zur Fehlererkennung vorgestellt, der erst nach Abschluss des Druckvorgangs durchgeführt wird. In [5] präsentieren die Autoren einen Ansatz zur Fehlererkennung, der auf einem speziell entwickelten binären Klassifikationsmodell eines faltenden neuronalen Netzes (engl. Convolutional Neural Networks, CNNs) basiert und eine Echtzeiterkennung von FFF-Defekten mit einer Genauigkeit von 84% ermöglicht. Allerdings ist dieser Ansatz aufgrund der Draufsicht-Kameraperspektive auf Füllmuster beschränkt, sodass das CNN-Modell keine weiteren Defektklassen erkennen kann. In [6] wurde ein Single-Shot-Detektor zur Erkennung von Fädenbildung in FFF-Prozessen vorgestellt. Zudem wird in [7] ein Ansatz zur Erkennung von Warping-Defekten beschrieben, bei dem Schlüsselpunkte zwischen dem gedruckten Produkt und einem Bild desselben verglichen werden. In [8] wurde ein modifizierter, leichterer Ansatz der vierten Generation der „You Only Look Once“ (YOLOv4)-CNN-Architektur zur Fehlererkennung entwickelt. Allerdings wurde in [9] ein homogener Datensatz zur Modellschulung verwendet, wodurch das YOLOv4-Modell nur eine bestimmte Art von FFF-Defekten erkennen kann. Zusammenfassend lässt sich festhalten, dass trotz der Bedeutung der Fehlererkennung in FFF-Prozessen bislang kein allgemeiner und automatisierter Ansatz existiert, der unabhängig von der Form des gedruckten Produkts und der Klasse des FFF-Defekts eine zuverlässige Erkennung ermöglicht.

In dieser Studie wird ein Ansatz zur automatisierten Fehlererkennung (engl. Automated Defect Detection, ADD) in FFF-Prozessen vorgestellt. Der ADD-Ansatz kombiniert tief lernende Modelle, insbesondere Klassifikationsmodelle CNN, mit Bildverarbeitungstechniken zur Erkennung von FFF-Defekten. Zudem berücksichtigt der ADD-Ansatz sechs äußere FFF-Defektklassen, nämlich Fädenbildung, Rissbildung, Materialansammlungen, sogenannte Spaghetti-Fehler, Unterextrusion und Schichtlücken.

2 Entwurf und Implementierung des ADD-Ansatzes

In diesem Abschnitt wird der Entwurf und die Implementierung des ADD-Ansatzes vorgestellt, der aus zwei Phasen besteht: (i) Datenvorbereitung und (ii) Entwicklung des Klassifikationsmodells, wobei jede Phase mehrere Schritte umfasst. Im Folgenden werden die beiden Phasen des ADD-Ansatzes kurz erläutert.

2.1 Phase 1: Datenvorbereitung

1- **Datenerfassung:** Ein Datensatz D_n, der aus n Bildern gedruckter Produkte besteht, wird aus öffentlich verfügbaren und offenen Quellen gesammelt. Der Datensatz D_n enthält Bilder sowohl von gedruckten Produkten mit FFF-Defekten, d. h. Fädenbildung, Rissbildung, Materialansammlungen, sogenannten Spaghetti-Fehlern, Unterextrusion und Schichtlücken, als auch von fehlerfrei gedruckten Produkten.

2- **Datenvorverarbeitung und -beschriftung:** Da die Bilder im Datensatz D_n aus verschiedenen Quellen stammen, können sie in Größe, Farbe und Betrachtungswinkel variieren. Um die Trainingsergebnisse in den nachfolgenden Schritten zu verbessern, werden die Bilder zunächst zugeschnitten und auf eine vordefinierte Größe skaliert. Anschließend werden die Bilder in Graustufen umgewandelt, um zu vermeiden, dass die Farben des FFF-Filaments unbeabsichtigt mit bestimmten FFF-Defektklassen korreliert werden. Schließlich werden die Bilder mit der entsprechenden FFF-Defektklasse beschriftet.

3- **Datenaugmentation und -aufteilung:** Um die Trainingsergebnisse des Modells zu verbessern und das Risiko von Underfitting oder Overfitting in der nächsten Phase zu minimieren, werden Techniken zur Datenaugmentation, wie Spiegeln, Rauschen, Belichtung („Helligkeit“), Drehen und zufällige Bildausschnitte, angewandt, um die Größe des Datensatzes D_n zu erhöhen. Anschließend werden die Bilder im Datensatz D_n in einen Trainingsdatensatz $D_{training}$, einen Validierungsdatensatz $D_{validation}$ und einen Testdatensatz $D_{testing}$ aufgeteilt.

2.2 Phase 2: Entwicklung des Klassifikationsmodells

1- **Architektur und Definition der Hyperparameter:** In diesem Schritt wird die Architektur des CNN definiert. Im Allgemeinen besteht ein CNN-Klassifikationsmodell aus einer Eingabeschicht, Faltungsschichten, Pooling-Schichten, voll verbundenen Schichten und einer Ausgabeschicht. Die Eingabeschicht verarbeitet und speist Bilder in die Faltungsschichten (auch als Kerne bekannt) ein, die Filter anwenden, um niedrigstufige Merkmale in den Bildern zu erkennen, wie Kanten oder Muster, die in Merkmalskarten organisiert sind. Pooling-Schichten reduzieren die Dimensionalität der Merkmalskarten, wodurch die Rechenlast verringert und die Merkmale robust gegenüber geringfügigen Variationen gemacht werden. Voll verbundene Schichten kombinieren und interpretieren hochgradige Merkmale, die von den Faltungs- und Pooling-Schichten extrahiert werden, um Klassifikationswerte zu erzeugen. Schließlich verwendet die Ausgabeschicht die Klassifikationswerte, um die Klassifikationskategorie für jedes Eingangsbild vorherzusagen, d. h. die FFF-Defektklasse.

2- **Modelltraining:** Nachdem die CNN-Architektur definiert wurde, wird der Trainingsdatensatz Dtraining verwendet, um das CNN-Klassifikationsmodell zu trainieren. Die Bilder in $D_{training}$ werden über die Eingabeschicht als 2D-Matrix $\mathbf{I}$ in die Modelle eingespeist, wobei jedes Iij die Graustufen-Pixelintensität an der Position ij ($0 < i, 0 < j$) im Bild darstellt, wie in Gleichung 1 gezeigt. Jeder Pixel hat einen Graustufen-Pixelintensitätswert, der von 0 bis 255 reicht (0

ist schwarz, 255 ist weiß, und Werte dazwischen repräsentieren Grautöne). Während des Trainings „lernt“ das Modell aus den bestehenden Beziehungen zwischen bekannten Eingaben (Bildern) und bekannten Ausgaben (FFF-Defektklassen, z. B. „Spaghetti“). Um während des Trainingsprozesses eine Überanpassung zu vermeiden, wird der Validierungsdatensatz $D_{validation}$ verwendet, um die Modelle zu optimieren und deren Leistung während des Trainings zu bewerten.

$$\mathbf{I} = \begin{bmatrix} I_{1,1} & \cdots & I_{1,j} \\ \vdots & \ddots & \vdots \\ I_{i,1} & \cdots & I_{i,j} \end{bmatrix}, \qquad I_{i,j} \in R \,|\, 0 \le I \le 255 \tag{1}$$

3- **Modellbewertung:** Nach Abschluss des Trainings wird der Testdatensatz $D_{testing}$ verwendet, um die Leistung der Klassifikationsmodelle zu bewerten. In diesem Schritt werden Bewertungsmetriken wie Präzision (pr), Rückruf (rc) und die mittlere durchschnittliche Präzision ($mAP50$) bei einem einzelnen Schnittstellenüberdeckungsgrad (IoU) von 50% verwendet. Einerseits zeigen pr und rc die Fähigkeit der Modelle an, richtig positive (TP) Klassifikationen zu erzielen und falsch positive (FP) sowie falsch negative (FN) Klassifikationen zu vermeiden. Die Gleichungen 2, 3 und 4 präsentieren die mathematischen Formulierungen von pr, rc und IoU, jeweils. CNN-Klassifikationsmodelle mit einem $mAP50$-Wert, der einen vordefinierten Schwellenwert überschreitet, werden für die Bereitstellung gespeichert. Andernfalls werden die Schritte 1 und 2 wiederholt, nachdem die CNN-Architektur modifiziert wurde. In Gleichung 4 ist Ω_o die überlappende Fläche zwischen dem vorhergesagten Bereich und der Ground Truth, und Ω_u ist die gesamte Fläche, die sowohl vom vorhergesagten Bereich als auch von der tatsächlichen Wahrheit abgedeckt wird.

$$pr = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{2}$$

$$rc = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{3}$$

$$IoU = \frac{\Omega_o}{\Omega_u} \tag{4}$$

3 Validierung des ADD-Ansatzes

Zur Validierung des ADD-Ansatzes wird ein Praxistest durchgeführt, bei dem gedruckte Produkte verwendet werden, die dazu entworfen wurden, FFF-Defektklassen nachzubilden. Das Hauptziel des Validierungstests besteht darin, die Fähigkeit des CNN-Klassifikationsmodells zu bewerten, die FFF-Defekte genau zu erkennen und folglich den 3D-Druckprozess zu stoppen. Für den Validierungstest wird ein Prusa Original i3 MK3S 3D-Drucker als Druckhardware eingesetzt. Ein Raspberry Pi 4 mit 8 GB RAM ist über ein USB-Kabel mit dem 3D-Drucker verbunden und fungiert als Steuereinheit zur Verwaltung des 3D-Druckers. Eine Logitech 1080p Pro Stream Webcam ist an

der z-Achse (vertikal) montiert, um den Druckbereich zu überwachen, und ist mit dem Raspberry Pi verbunden. Abbildung 1 zeigt das Hardware-Setup des Validierungstests.

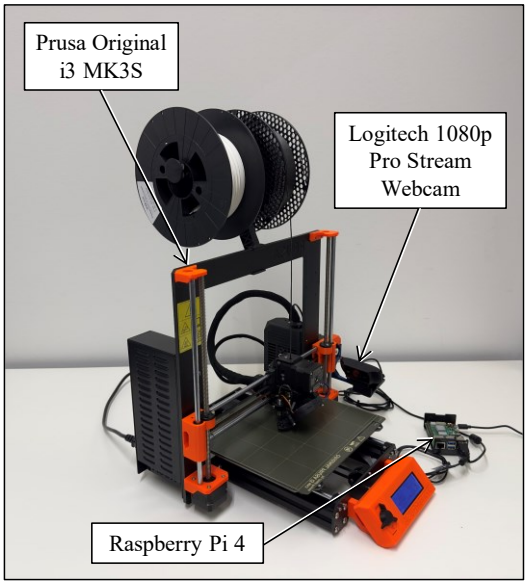


Abbildung 1: Setup des Validierungstests

Für die Datenvorbereitung wird ein Datensatz der Größe $n = 4{,}124$ gesammelt ($D_{n=4,124}$) aus öffentlich zugänglichen und offenen Quellen wie Roboflow Universe, r/FixMyPrint und r/3Dprinting auf Reddit sowie YouTube, der Bilder von 3D-gedruckten Produkten mit FFF-Defekten sowie fehlerfreie Bilder umfasst. Im $D_{n=4,124}$ enthalten insgesamt 832 Bilder 3D-gedruckte Produkte mit FFF-Defekten und $3{,}292$ fehlerfreie Bilder. Anschließend werden die Bilder im Datensatz nacheinander in die ADD-Vorverarbeitungsanwendung importiert und auf eine Größe von $640{\times}640$ Pixeln skaliert, was einen Kompromiss zwischen effizientem Batch-Training innerhalb der Systemgrenzen und der Erhaltung ausreichender Bilddetails darstellt. Die Datenaugmentation, die mit der ADD-Vorverarbeitungsanwendung durchgeführt wird, erfolgt durch das Erzeugen von zwei Kopien jedes Bildes im $D_{n=4,124}$, von denen eine gespiegelt wird. Danach werden die beiden Kopien jedes Bildes einem Rauschen ausgesetzt, bei dem 4% des Graustufen-Pixelintensitätswerts I_{ij} an den Positionen i und j $(0 < i < 640, 0 < j < 640)$ durch einen zufälligen Wert zwischen 0 und 255 ersetzt werden (siehe Phase 2, Schritt 2). Anschließend werden die beiden Kopien einer zufälligen Helligkeitsänderung von ±35% sowie einer zufälligen Drehung von ±20° ausgesetzt. Infolgedessen wird der Datensatz $D_{n=4,124}$ auf $n = 12{,}372$ erweitert und in einen Trainingsdatensatz (80%) $D_{training=9,897}$, einen Validierungsdatensatz (15%) $D_{validation=1,855}$ und einen Testdatensatz (5%) $D_{testing=620}$ aufgeteilt.

Die CNN-Architektur wird ausgewählt, nachdem vorläufige Tests mit vier YOLO-Architekturen durchgeführt wurden, nämlich der YOLOv8l, YOLOv9c, YOLOv10l und YOLOv10b. Die Anzahl der Trainingsepochen wird auf 250 festgelegt, und der Verlust wird in jeder Epoche berechnet. Der Verlust in YOLO basiert auf mehreren Komponenten, darunter: (i) Boxverlust für die Lokalisierung von Begrenzungsrahmen, (ii) binärer Kreuzentropieverlust für die Klassifizierung (Klassenverlust)

und (iii) dualer Fokalverlust zur Verfeinerung der Begrenzungsrahmen, um eine genaue und robuste Objekterkennung zu gewährleisten [9]. Periodisch (nach Abschluss einer bestimmten Anzahl von Epochen) wird der Validierungsdatensatz $D_{validation=1,855}$ durch das CNN propagiert, um dessen Verlust („Validierungsverlust“) zu bewerten und sicherzustellen, dass der Validierungsverlust denselben Trend wie der Trainingsverlust folgt, um eine Überanpassung zu vermeiden. Das Training wird für alle Kandidaten-CNN-Architekturen durchgeführt; Abbildung 9 zeigt die Verlustmetriken für $D_{training=9,897}$ der YOLOv9c.

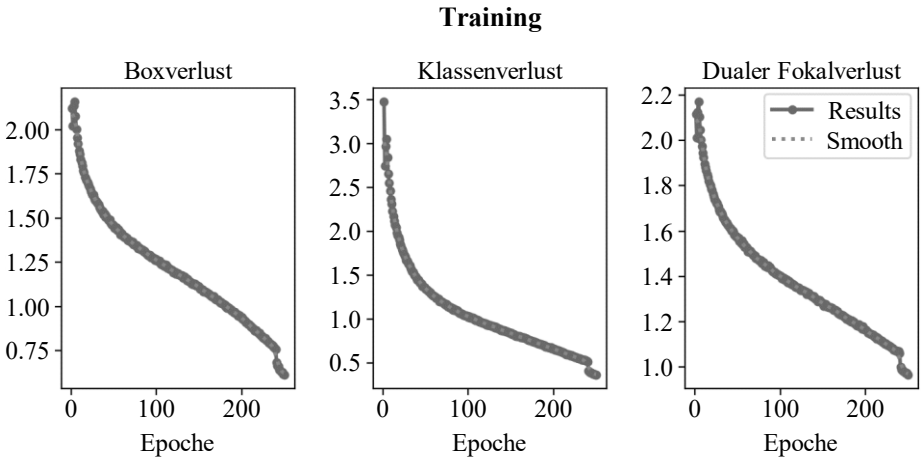


Abbildung 2: Verlustmetriken $D_{training=9,897}$ im Vergleich zu den Trainingsepochen für das auf der YOLOv9c-Architektur basierende Klassifikationsmodell

Die Leistung der Kandidaten-CNN-Architekturen wird anhand des Datensatzes Dtesting= 620 bewertet, der während des Trainings nicht eingeführt wurde. Tabelle 1 umfasst die Leistung aller Kandidatenarchitekturen hinsichtlich der Verlustmetriken und *mAP50*-Werte. Infolgedessen wird die Architektur YOLOv9c für das CNN-Klassifikationsmodell in dieser Studie ausgewählt, da sie den höchsten *mAP50*-Wert von 82,2% aufweist.

Table 1: Sample table with three columns and five rows

Modellarchitektur	**Boxverlust**	**Klassenverlust**	***mAP50 (%)***
YOLOv8l	0.824	2.101	37.6%
YOLOv9c	1.497	0.890	82.2%
YOLOv10l	3.118	1.978	81.7%
YOLOv10b	3.107	2.058	81.7%

Als nächstes wird ein CAD-Modell entworfen, das zwei vertikale Säulen auf einer Grundplatte zeigt. Für den Druck wird ein Filament aus Polymilchsäure (PLA) verwendet, dessen Düsentemperatur typischerweise im Bereich von 180-220°C liegt. Um absichtlich einen FFF-Defekt, insbesondere Fädenbildung, zu induzieren, wird die Düsentemperatur auf 260°C eingestellt.

4 Ergebnisse und Diskussion

Die Ergebnisse des Validierungstests, die in Abbildung 3a dargestellt sind, bestätigen die Effektivität des ADD-Ansatzes bei der Erkennung von FFF-Defekten auf einem Nachweisniveau. Das auf der YOLOv9c-Architektur basierende CNN-Klassifikationsmodell identifizierte erfolgreich den Fädenbildungsdefekt und pausierte automatisch den 3D-Drucker mithilfe des Pausalgorithmus. Dadurch konnten Zeit- und Materialverschwendung reduziert und die Notwendigkeit einer menschlichen Überwachung eliminiert werden. Insbesondere erreichte der Fädenbildungsdefekt die zuvor festgelegte Persistenzzeitschwelle von 20 Sekunden und führte somit zur Unterbrechung des FFF-Prozesses. Darüber hinaus ist in Abbildung 3b zu erkennen, dass Fehlklassifikationen des CNN-Modells auf ein Minimum reduziert werden konnten. Konkret wurden keine Identifizierungen von Materialansammlungen, Spaghetti-Fehlern, Unterextrusion oder Schichtlücken beobachtet. Unter bestimmten Lichtbedingungen klassifiziert das CNN-Klassifikationsmodell jedoch Teile des 3D-Druckers, insbesondere den Extruder, fälschlicherweise als Rissbildungsdefekt. Ein ähnliches Problem tritt auch bei den Gitterlinien des Druckbetts auf, bei denen die Gitterlinien als Rissbildungsdefekte erkannt werden.

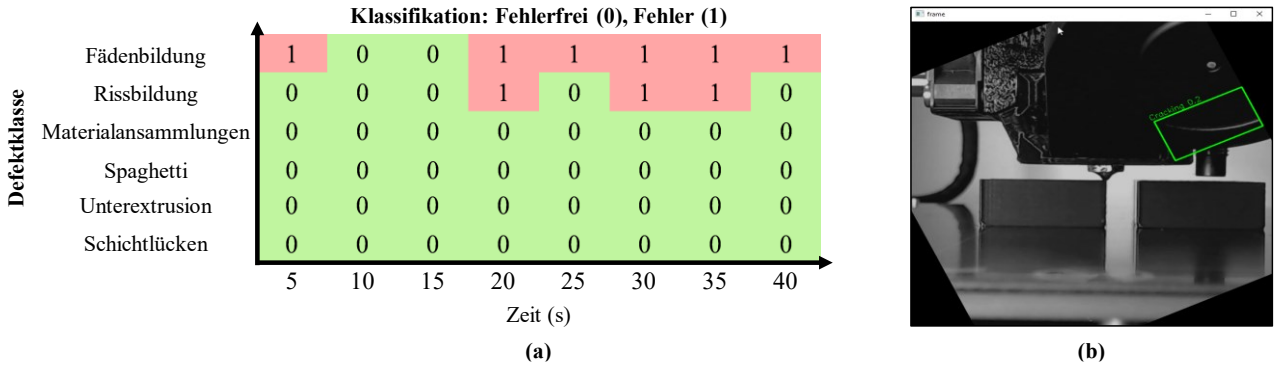


Abbildung 3: (a) Ereigniswarteschlange und Fehlererkennung im Validierungstest, (b) Fehlklassifizierung des Extruders des 3D-Druckers als Rissbildungsdefekt unter bestimmten Lichtbedingungen

5 Zusammenfassung und Schlussfolgerungen

In dieser Arbeit wurde ein automatisierter Ansatz zur Fehlererkennung in FFF-Prozessen vorgestellt, der Deep Learning und Bildverarbeitung kombiniert. Der ADD-Ansatz vereint CNNs und Bildverarbeitung, um Defekte in FFF-Prozessen autonom zu erkennen. Validierungstests wurden durchgeführt, zunächst durch die Definition der CNN-Architektur, basierend auf vorläufigen Tests mit vier Kandidaten-Architekturen: YOLOv8l, YOLOv9c, YOLOv10l und YOLOv10b. Anschließend wurde das CNN-Klassifikationsmodell auf Basis der YOLOv9c-Architektur ausgewählt, da es die beste Leistung mit einem $mAP50$-Wert von 82,2% aufwies. Die Bilder der gedruckten Produkte wurden gezielt entworfen, um die Klassen von FFF-Defekten zu simulieren. Die Testergebnisse haben die Fähigkeit des ADD-Ansatzes zur Erkennung von Defekten in FFF-Prozessen auf einem

Nachweisniveau bewiesen, indem ein Fädenbildungsdefekt korrekt identifiziert und der Druckprozess angehalten wurde. Dadurch konnten Zeit- und Materialverschwendung sowie die Notwendigkeit einer menschlichen Überwachung in FFF-Prozessen reduziert werden.

Danksagungen

Die Autoren möchten sich herzlich für die Unterstützung durch das Bundesministerium für Digitales und Verkehr (BMDV) im Rahmen des mFUND-Programms unter den Förderkennzeichen 01FV2013B und 01FV2059C sowie durch die Deutsche Forschungsgemeinschaft (DFG) unter dem Förderkennzeichen SM 281/31-1 bedanken. Alle in dieser Arbeit geäußerten Meinungen, Erkenntnisse, Schlussfolgerungen oder Empfehlungen sind die der Autoren und spiegeln nicht notwendigerweise die Ansichten des BMDV oder der DFG wider.

Literaturverzeichnis

[1] Dragos, K. and Smarsly, K., 2023. Damage detection coupling convolutional neural networks and numerical simulations. In: Proceedings of the 30th International Workshop on Intelligent Computing in Engineering (EG-ICE). London, United Kingdom, 07/04/2023.

[2] Dadoulis, G., Manolis, G. D., Katakalos, K., Al-Zuriqat, T., Dragos, K. and Smarsly, K., 2024 Towards detecting damage in lightweight bridges with traveling masses using machine learning. In: Proceedings of the 2024 European Conference on Computing in Construction (EC3). Chania, Crete, Greece, 07/15/2024.

[3] Al-Zuriqat, T., Chillón Geck, C., Dragos, K., & Smarsly, K., 2023. Adaptive fault diagnosis for simultaneous sensor faults in structural health monitoring systems. Infrastructures, 8(3), 39.

[4] Zhao, X., Li, Q., Xiao, M., and He, Z., 2023. Defect detection of 3D printing surface based on geometric local domain features. The International Journal of Advanced Manufacturing Technology, 125(1), pp.183 – 194.

[5] Khan, M.F., Alam, A., Siddiqui, M.A., Alam, M.S., Rafat, Y., Salik, N., and Al-Saidan, I., 2021. Real-time defect detection in 3D printing using machine learning. Materials Today: Proceedings, 42(2021), pp.521 – 528.

[6] Paraskevoudis, K., Karayannis, P., and Koumoulos, E.P., 2020. Real-time 3D printing remote defect detection (stringing) with computer vision and artificial intelligence. Processes, 8(11), p.1464.

[7] Xie, J., Saluja, A., Rahimizadeh, A., and Fayazbakhsh, K., 2022. Development of automated feature extraction and convolutional neural network optimization for real-time warping monitoring in 3D printing. International Journal of Computer Integrated Manufacturing, 35(8), pp.813-830.

[8] Xu, L., Zhang, X., Ma, F., Chang, G., Zhang, C., Li, J., Wang, S., and Huang, Y., 2023. Detecting defects in fused deposition modeling based on improved YOLO v4. Materials Research Express, 10(9), p.095304.

[9] Ultralytics (n.d.-a) Ultralytics Models Documentation. Ultralytics. Available at: https://docs.ultralytics.com/models/ (Accessed: 19 October 2024).

Concept for digitally AI-readable buildings for urban mining based on digital building logbooks using large language models

Marcel Heiß[1], Benedikt Kandler[1] and Uwe Rüppel[1]
[1]Institute of Numerical Methods and Informatics in Civil Engineering, Darmstadt, Germany
E-mail(s): heiss@iib.tu-darmstadt.de

Abstract: The construction sector holds great potential for contributing to climate neutrality through circular economy strategies such as Urban Mining. But reuse of building materials often is hindered by a lack of structured, machine readable information on the existing building stock. This paper presents a concept for enhancing Digital Building Logbooks (DBL's) through the combined use of Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG). By semantically processing unstructured documents such as reports and invoices, the proposed approach enables AI-readable DBL's that facilitate data-driven decision-making for Urban Mining. The study analyses DBL's and current barriers like lack of accessible and interoperable data sources and static nature of DBL and counters them with introducing a system architecture that transforms unstructured building documentation into a dynamic DBL. This AI-enabled DBL approach significantly enhances data accessibility, supports material inventories and fosters reuse scenarios in the built environment.

Keywords: Large Language Models, Urban Mining, Circular economy, Digital Building Logbook

DOI: 10.18154/RWTH-CONV-254913. Published in the conference proceedings of the 36. Forum Bauinformatik 2025, Aachen, Germany.
© 2025 The copyright for this article lies with the authors. This publication, except for quotations and otherwise indicated parts, is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

1 Introduction

The construction sector is responsible for a considerable portion of resource consumption and waste generation, and also emits significant levels of greenhouse gases [1], [2]. Concurrently, the built environment constitutes society's largest resource stock. Urban Mining offers a pathway to reduce the high consumption and dependency on virgin materials, by recovering and reusing resources, as well as integrating them into a circular economy (CE). A major barrier is the lack of information required. For CE and urban mining, creating a material inventory of individual buildings and representing it digitally at city level through a GIS-based material cadastre is the first step towards knowledge-based decision-making, followed by digital building material passports and digital product passports, which quantify the sustainability of buildings, components and materials. Currently, extensive on-site recording and manual information extraction from documents are necessary for existing buildings to obtain this information and create detailed digital building material passports. Consequently, the idea of a whole-life-cycle repository of building information in the form of DBL's is becoming increasingly

important in order to gather and consolidate building data. Especially documentation of existing buildings often consists of unstructured text documents. These documents are part of the current state of DBL development, but are not analysed further. Concurrently, the performance of LLM's is witnessing a significant enhancement by handling unstructured text data. This research proposes a framework for using LLM in RAG to handle unstructered text data in DBL environment.
This work comprises a conceptual study beginning with the current state of development of DBL and LLM. The following consolidation analyses the shortcomings of current DBL implementations and examines their approaches to integrating unstructured, heterogeneous data sources to build DBL's, highlighting the research gap and its potential. Building on this, a DBL framework is developed that applies RAG and LLMs to extract structured data from unstructured documents and transfers gathered data into a vector that represents the "building construction DNA". After that the approach is analysed in a discussion with regard to solving the identified deficits and its potential for urban mining.

2 Related Work

2.1 Digital Building Logbook

A DBL is defined as a centralised and common repository of building-related information in digital form that can be used for effective decision making throughout the life cycle of a building [3]. It dynamically collects, stores and makes available information on a variety of aspects of a building, including design, construction, quality and environmental and performance criteria. The continuous collection of data provides a basis of information on the current status as well as the history and thus offers transparency and accessibility to this information for different stakeholders compared to traditional paper-based building records. Synonyms include 'property logbook', 'building passport' and 'electronic building file', but it is distinct from the terms 'digital building material' and 'building resource passport', which only contain essential circularity information for each individual building. The research projects European Union Building SuperHub (EUB SuperHub) defined a DBL data structure and thus laid the foundation for the general use of DBL's [4]. To determine the characteristics of a building, the main categories of DBL containing individual properties are defined. Examples of central main categories are general and administrative information, building performance and building operation and usage and can be extended by main categories like Smart Readiness and Finance [4]. Current Research has focused primarily on the application of DBL's in the context of renovation planning and energy efficiency to achieve sustainability [4], [5]. To achieve climate neutrality, the focus is also shifting to CE in the construction industry and is the subject of research in the Demo-Blog project [5]. While current DBL approaches are based on machine readable data, which are gained from existing building data sources or user inputs and follow a determined data flow [6], a large part of the information is currently still recorded in the form of documents and is therefore not available for current approaches. Recent developments in LLMs provide new methods to extract and utilise information from unstructured data.

2.2 Large Language Models

LLMs represent a paradigm shift in natural language processing (NLP) driven by the transformer architecture, increased computational capabilities and the availability of large-scale training data [7]. It represents a type of artificial intelligence model trained on massive amounts of text data to perform

various natural language processing tasks. The performance of LLMs correlates strongly with the number of parameters, the amount of data and the computing power [8]. Performance for specific applications can be increased through pretraining and fine-tuning. While pretraining is usually carried out in autoregressive mode on huge amounts of text, fine-tuning can be general or domain-specific. Despite these methods, LLMs face limitations in their reasoning abilities. In so-called hallucinations, models can generate plausible-sounding but false statements [9]. Limiting the number of tokens that can be processed results in a limited context scope and often the training data is not up-to-date so that information after the training period is unknown. This is why approaches such as RAG were developed. RAG combines LLMs with external knowledge sources through embedding search and dynamic information integration [10]. First, the text documents and the query are vectorised using an embedding encoder and stored in a vector database. In the initial retrieval step, relevant documents are extracted from a knowledge database via vector search. In the augmented step, the content found is transferred to the LLM together with the original query. The LLM then generates a response based on the updated, fact-based information in selected documents.

3 Consolidation of DBL and combination with AI

Previous approaches of DBL are mainly concerned with the content and structure of DBL's but don't define a procedure to automatically gather, store and share information about the lifecycle of a building [11]. Literature identifies national or regional databases, BIM models, building monitoring-related sources and European initiatives as the four main categories potential DBL data sources. Building documents often part of a DBL in the form of a document repository [12]. Current governmental data relevant to DBLs is often incomplete, not interoperable, and not openly accessible [6]. Core sources such as cadastral data provide basic building attributes but seldom contain information relevant to urban mining, such as material composition [6] . However, this information can be derived from detailed descriptions of structures contained in building performance and thermal insulation certifications, as well as from plans from which quantities can be extracted. In conclusion, more information needs to be collected on the actual condition of the buildings [6]. For new buildings, BIM facilitates the creation of machine-readable data, but such information frequently remains siloed within specialist disciplines and is rarely transferred in a structured form into the operational phase. [11]. For existing buildings, full digital capture (e.g., as-built BIM) is uncommon, further limiting DBL digitisation levels. The evaluation of different initiatives with regard to their online accessibility and level of digitisation showed that the majority of these only represent paper-based documents and only a few approaches allow a digital twin integration and the fewest allow an automatic input of data for example from BIM model [13].
The first part of the analysis shows current deficits of DBL in the form of a lack of data sources and their lack of accessibility, interoperability and low level of digitisation. This results in barriers to using DBL, including the static nature of DBL, as information often has to be updated manually, and the limited accessibility for various stakeholders [13]. Unstructured and heterogeneous documents are not characterised as data sources. Some initiatives, such as OpenDBL and DigiBuild, address heterogeneous data integration through ontologies, ETL processes, and microservice architectures [14], [15]. Hybrid approaches like the EUB SuperHub link distributed data sources via building UUIDs,

while frameworks like Demo-BLog enrich stored files with metadata to support future processing [4], [16]. Included CLEA system uses ETL to extract data from national repositories and a microservice architecture to integrate sensor data [17]. However, across all approaches, unstructured documents are generally stored but not analysed. The absence of methods for extracting meaningful, structured information from such data represents a critical research gap, which recent advances in LLMs could address [12].

4 Adapted DBL framework for Urban mining using LLM

The analysed research shows different approaches to interoperate different heterogeneous data sources to create a database for DBL using different digital technologies. Based on [6], [18], the system model in figure 1 summarises the main findings in a structure of using different data sources for DBL with the help of digital methods. In addition to the inputs of the DBL and their preprocessing to ensure their use in the DBL, a processing and output sequence is added, which represents the use of DBL data for various use cases.

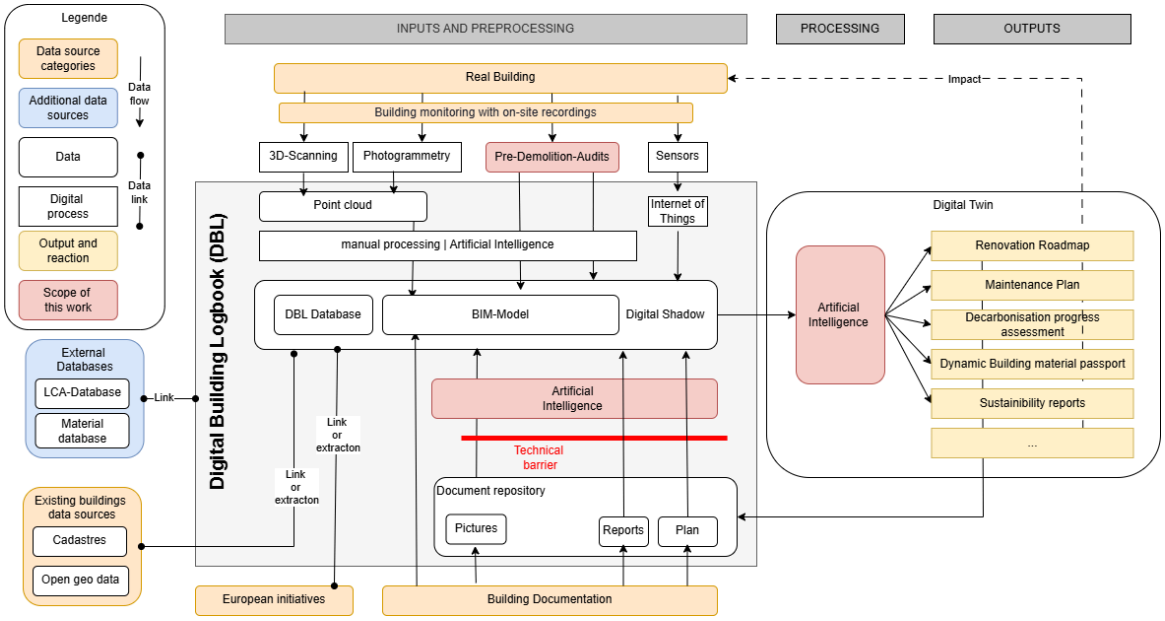


Figure 1: Scope of work and problem definition in context of DBL

"Inputs and preprocessing" is based on four categories of data sources: existing buildings, European initiatives, building documentation and building monitoring. Existing databases, such as cadastres, land registries or open geo data, can be included in DBL as an initial data set by extracting them, for example, using ETL, or by making them interoperable via links and ontologies [6], [15]. In addition to existing databases, new initiatives currently being developed, such as the Digital Product Passport, can be made accessible through extraction and linking. Another data source is building monitoring with on-site recordings. The most established approaches are those for the geometric recording of buildings using 3D scanning and photogrammetry, which can be further processed manually or by AI to generate a BIM model. Furthermore, the use of sensors and Internet of Things (IoT) technology is widespread during the use phase of buildings to capture various time series data for supervising and

controlling a building by coupling it with a BIM model to create a digital shadow of the building. For Urban Mining and the circular use of materials and components pre-demolitions audits (PDA) are of crucial importance to create an information basis.

While building documentation is often captured mainly in reports and plans, supplemented by images, and remains within a document repository due to the absence of technical methods, BIM models can be used directly as a data source in DBL. The lack of integration is a major issue and represents a crucial research gap. The utilisation of the various data sources and their pre-processing result in the creation of a digital shadow of the building, which provides data in DBL database, links, BIM model and other forms. The ultimate goal is to use the data obtained as an information basis for various questions in a processing step, which often generates outputs in the form of reports. At building level, for example, building material passports and renovations roadmaps play a decisive role in assessing sustainability and characterising measures to improve it. In addition to integrating the outputs into the document repository of the DBL, the impact in the form of the measures to be implemented on the real building plays a key role.

The research question explores how the use of LLMs can overcome this technical barrier and render this unstructured information usable in the context of DBL and further using AI in processing step.

4.1 Adapted RAG dataflow structure

The proposed approach involves adapting the technical architecture model for RAG to create a dynamic DBL for semantic indexing and utilisation of unstructured documents. The aim is to extract specific DBL information from different unstructured documents and transfer it into an overview. Further information is queried from the user via prompts and answered according to the RAG principle, with reference to the relevant sources.

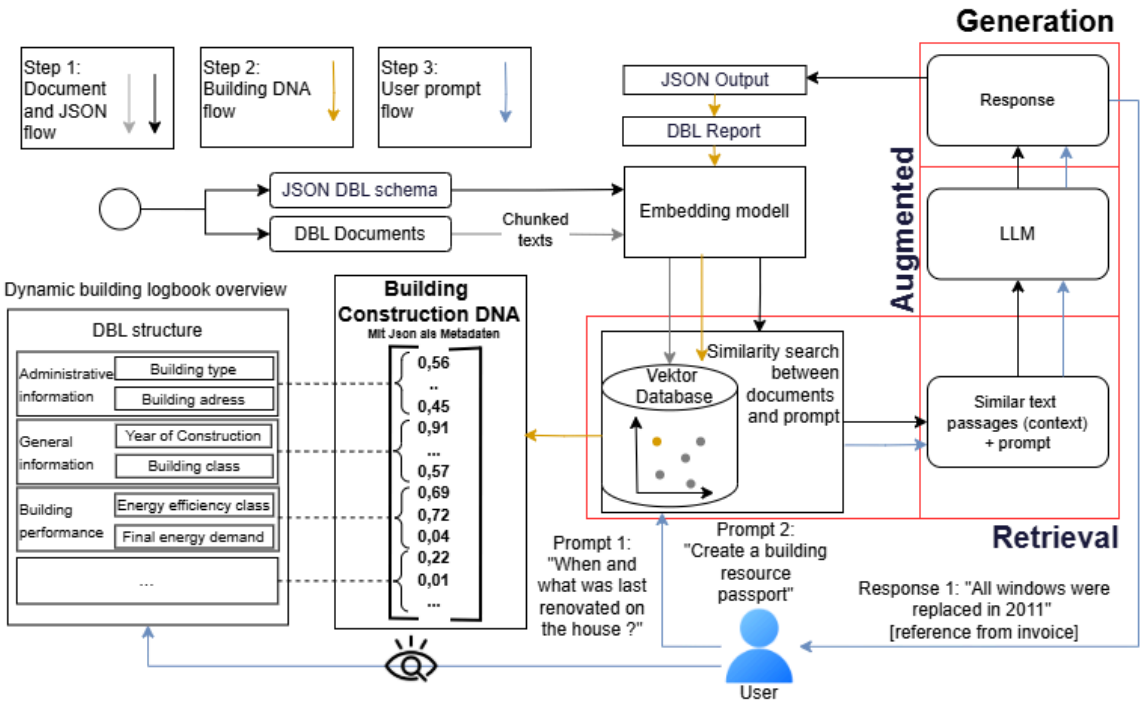


Figure 2: Adapted RAG dataflow structure

The starting point is the unstructured documents within the document repository and a JSON Schema, which defines the expected data to be included in the DBL and extracted from the documents. The workflow begins with document pre-processing. All building-related documents—reports, plans, certificates—are split into semantically coherent text chunks. Each chunk is converted into a high-dimensional vector using an embedding model. Alongside the vector, metadata such as the source document, section position, and chunk length is stored. These embeddings are kept in a vector database optimised for similarity search.
In the retrieval process of RAG vector database is searched for semantically closest chunks to the JSON schema .The relevant excerpts from the external knowledge base in form of DBL documents are used in the Augmented step for a context-enriched prompt that serves as input for the LLM, which generates a response based on it. This response is then transferred to the JSON output format, which structures the information according to the JSON DBL schema.
From this structured output the second step then begins, aiming to create a vector-based building construction DNA. Technically, this DNA is a second-level embedding that encodes the complete set of structured building attributes into a single high-dimensional vector. This representation functions as a compressed digital fingerprint of the building. First, the JSON output is converted into a text-based report by an LLM and then converted into a vector representation by the embedding model and stored in a vector database. Referring to human DNA, the construction building DNA represents the encoded blueprint for the building and contains all the information on how the building functions. This is then used to generate a dynamic building logbook overview in the form of a structured report, which bundles all the information from the DBL previously defined in the JSON schema. The third step deals with user interaction. Users can view the dynamic building logbook overview to gain quick insight into the building's key figures. Further information can be entered via a chatbot function in the form of prompts, based on the RAG principle and the document vector representations.
In practice, the Building Construction DNA serves as an indexable and comparable representation of a building's technical state within the DBL system. By storing both chunk-level embeddings (fine-grained retrieval) and DNA-level embeddings (holistic building profile), the framework supports both targeted queries and high-level building comparisons.

5 Discussion

The presented approach addresses the limited number of DBL data sources by making unstructured documents accessible as data sources. Despite the low level of digitisation of the data sources, the approach achieves a high level of digitisation. This is achieved by universally storing the data in a vector database and accessing it via a chatbot and LLMs. The continuous integration of new documents results in a dynamic DBL, thus releasing a static structure and the need for manual updates of information. The approach is limited by the use of non-open data, so the voluntary participation of those involved is needed. Technical issues also need to be clarified, such as how to handle contradictory information in documents on the same topic, and how to manage document versions. Further research is needed to establish the interoperability of this approach with other data sources. The approach has further potential with other unstructured data sources, such as plans and images,

as well as structured data sources, such as databases. This approach offers two significant potential benefits for urban mining. Firstly, the inclusion of PDA enables the integration of specialised data, such as the higher-value utilisation paths of components, in an unstructured form. Secondly, the use of this and all other data to generate material and component inventories in the dynamic DBL provides a decisive basis for urban mining. In addition, the vector database created can also serve as a basis for building material cadastres on city level.

6 Conclusion

The presented concept demonstrates how LLMs can overcome the current limitations of DBL, such as the lack of data sources and the inability to handle unstructured data. By embedding document content and integrating similarity-based retrieval with natural language generation, the system can transform static document repositories into dynamic knowledge bases. However, other challenges remain, such as data privacy and validation, particularly when handling contradictory information. Additionally, the approach relies on the voluntary participation of building owners who possess the relevant documents. However, this approach enables the digitisation of DBLs, particularly for existing buildings, and ensures that data relevant to urban mining and the circular economy can be automatically extracted, structured and utilised.

References

[1] G. A. f. B. United Nations Environment Programme and Construction, *Not just another brick in the wall: The solutions exist - Scaling them will build on progress and cut emissions fast. Global Status Report for Buildings and Construction 2024/2025*, Mar. 2025.

[2] M. de Wit, J. Hoogzaad, S. Rumjumar, H. Friedl, and A. Douma, “The Circularity Gap Report”, Circle Economy, Amsterdam, 2019.

[3] E. Commission, E. A. for Small, M.-s. Enterprises, *et al.*, *Definition of the Digital Building Logbook – Report 1 of the Study on the Development of a European Union Framework for Buildings' Digital Logbook*. Publications Office of the European Union, 2020. DOI: doi/10.2826/480977.

[4] M. Malinovec Puček, A. Khoja, E. Bazzan, and P. Gyuris, “A Data Structure for Digital Building Logbooks: Achieving Energy Efficiency, Sustainability, and Smartness in Buildings across the EU”, *Buildings*, vol. 13, no. 4, p. 1082, Apr. 20, 2023, ISSN: 2075-5309. DOI: 10.3390/buildings13041082.

[5] J. D. S. Gonçalves, W. C. Lam, and M. Ritzen, “The Role of Digital Building Logbooks for a Circular Built Environment”, in *A Circular Built Environment in the Digital Age*, C. De Wolf, S. Çetin, and N. M. P. Bocken, editors, Cham: Springer International Publishing, 2024, pp. 229–243, ISBN: 978-3-031-39674-8 978-3-031-39675-5. DOI: 10.1007/978-3-031-39675-5_13.

[6] M. Gómez-Gil, M. M. Sesana, G. Salvalai, A. Espinosa-Fernández, and B. López-Mesa, “The Digital Building Logbook as a gateway linked to existing national data sources: The cases of Spain and Italy”, *Journal of Building Engineering*, vol. 63, p. 105461, Jan. 2023, ISSN: 23527102. DOI: 10.1016/j.jobe.2022.105461.

[7] A. Vaswani, N. Shazeer, N. Parmar, *et al.*, “Attention Is All You Need”, in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017. DOI: 10.48550/ARXIV.1706.03762.

[8] J. Kaplan, S. McCandlish, T. Henighan, *et al.* “Scaling Laws for Neural Language Models”. arXiv: 2001.08361 [cs]. (Jan. 23, 2020), [Online]. Available: http://arxiv.org/abs/2001.08361 (visited on 04/16/2025), pre-published.

[9] Z. Ji, N. Lee, R. Frieske, *et al.*, “Survey of Hallucination in Natural Language Generation”, version 7, *ACM Computing Surveys*, vol. Volume 55, 2022. DOI: 10.1145/3571730.

[10] P. Lewis, E. Perez, A. Piktus, *et al.* “Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks”. version 4. (2020), [Online]. Available: https://arxiv.org/abs/2005.11401 (visited on 04/16/2025), pre-published.

[11] M. Gómez-Gil, A. Espinosa-Fernández, and B. López-Mesa, “Review and Analysis of Models for a European Digital Building Logbook”, *Energies*, vol. 15, no. 6, p. 1994, Mar. 9, 2022, ISSN: 1996-1073. DOI: 10.3390/en15061994.

[12] R. Alonso, R. Olivadese, A. Ibba, and D. Reforgiato Recupero, “Towards the definition of a European Digital Building Logbook: A survey”, *Heliyon*, vol. 9, no. 9, e19285, Sep. 2023, ISSN: 24058440. DOI: 10.1016/j.heliyon.2023.e19285.

[13] Executive Agency for Small and Medium sized Enterprises., *Building Logbook State of Play: Report 2 of the Study on the Development of a European Union Framework for Buildings’ Digital Logbook.* LU: Publications Office, 2020. [Online]. Available: https://data.europa.eu/doi/10.2826/519144 (visited on 06/16/2025).

[14] “openDBL Project. D1.3 System Specifications and AI-based mapping concept”. (2023), [Online]. Available: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e501f4f81d&appId=PPGMS.

[15] J. L. Hernández, D. Arévalo, S. Martín, *et al.*, “Using Extraction, Transformation and Loading Procedures for Digitalisation of Buildings”, in *2024 9th International Conference on Smart and Sustainable Technologies (SpliTech)*, Bol and Split, Croatia: IEEE, Jun. 25, 2024, pp. 1–6, ISBN: 978-953-290-135-1. DOI: 10.23919/SpliTech61897.2024.10612635.

[16] A. Ibba, S. A. Hwang, D. R. Recupero, R. Alonso, and R. Olivadese, “Defining Valuable Data and Stakeholder Engagement in Digital Building Logbooks: A Framework for Effective Data Storage”, presented at the 2024 European Conference on Computing in Construction, Jul. 14, 2024. DOI: 10.35490/EC3.2024.207.

[17] Redmond, Alan Martin, “The conceptual architecture requirements for french digital building logbook”, in *IARIA Congress 2024*, Porto, Portugal, 2024, pp. 81–87, ISBN: 978-1-68558-180-0.

[18] M. Gómez-Gil, A. Espinosa-Fernández, and B. López-Mesa, “Contribution of New Digital Technologies to the Digital Building Logbook”, *Buildings*, vol. 12, no. 12, p. 2129, Dec. 4, 2022, ISSN: 2075-5309. DOI: 10.3390/buildings12122129.

FORUMBAUINFORMATIK

gia | RWTH AACHEN UNIVERSITY

Design Computation | RWTH AACHEN UNIVERSITY

e3D | RWTH AACHEN UNIVERSITY

© Peter Winandy