Proceedings of the Hamburg International Conference of Logistics (HICL) – 27

Hendrik Haße, Bin Li, Norbet Weißenberg, Jan Cirullies and Boris Otto

Digital Twin for Real-Time Data Processing in Logistics

Published in: Artificial Intelligence and Digital Transformation in Supply Chain Management
Wolfgang Kersten, Thorsten Blecker and Christian M. Ringle (Eds.)
September 2019, epubli

Digital Twin for Real-Time Data Processing in Logistics

Hendrik Haße[1], Bin Li[1], Norbet Weißenberg[1], Jan Cirullies[1] and Boris Otto[1]

1 – Fraunhofer Institute for Software and Systems Engineering ISST

Purpose: Key performance indicators (KPIs) are an essential management tool. Real-time KPIs for production and logistics form the basis for flexible and adaptive production systems. These indicators unfold their full potential if they are seamlessly integrated into the "Digital Twin" of a company for data analytics.

Methodology: We apply the Design Science Research Methodology for Information Systems Research for deriving a digital twin architecture.

Findings: Research in the field of digital twins is at an early state, where the main objective is to find new applications for this technology. The majority of digital twin applications relate to the fields of manufacturing. Finally, it became apparent that existing architectures are too generic for usage in logistics.

Originality: The approach presented is an affordable solution for stakeholders to start with a digital transformation, based on standards and therefore highly technology-independent. The combined use of a lambda architecture with a semantic layer for flexible KPI definition is a special case.

Keywords: Digital Twin, Real-time, KPI, IoT

First received: 19.May.2019 **Revised:** 28.May.2019 **Accepted:** 11.June.2019

1 Introduction

Every day, logistics generates a vast amount of data, which is mainly generated by controlling and monitoring enormous flows of goods (Jeske, Grüner and Weiß 2014, p. 9). The data generated in this way holds considerable potential for optimization. A central challenge is the intelligent use of data (Spangenberg, et al., 2017, p. 44). The value of data is not measured by the amount of data collected, but by the applications made possible by the data. For this purpose, the collected data must be prepared in such a way that it can form the basis for optimization measures.

Making use of such data requires a substantial and valid data basis. Data collection, for example, is no longer a particular challenge due to increasingly improved and cheaper sensor technology. What is essential, however, is how this data is evaluated and how this evaluated data contributes to improving the specific process. Another important aspect is the processing time required to evaluate the collected data. The processing of large amounts of data, such as that generated by IoT (Internet of Things) applications, requires a particular framework in order to evaluate these enormous amounts of data (Mishra, Lin and Chang 2014, p. 124). Enterprises have to cope with an ever-increasing amount of data, which be-comes increasingly more efficient with the use of big data frameworks (Gupta, et al. 2017, p. 9).

Hence, it is generally no coincidence that the field of big data analytics plays such an important role in logistics. Logistics, with its cross-sectional function, is a key success factor, making big data analytics increasingly a strategic tool (Spangenberg, et al., 2017, p. 44; Hazen, et al. 2016, p. 592). The determination of key performance indicators (KPIs) is an essential management tool that allows a variety of different evaluations and analyses (Chae 2009, p. 423). This form of data processing and data visualization is made possible by the digital representation of physical assets in the form of a digital twin.

Particularly in logistics, the use of real-time data is an important instrument for visualizing events immediately (Park et al. 2014, 5). However, an industrial application of digital twin frameworks can be found mainly in the context of product management, shop floor and production management (Zhuong, et al. 2018, p. 1153; Qi, et al. 2018, p. 238). In addition, Hopkins and Hawking point out that there is a lack of real life use cases in logistics for both IoT applications and big data analytics (Hopkins and Hawking 2018, p. 579).

Motivated by these aspects, the approach chosen in this paper is to develop a data processing architecture that is tailored to the needs of logistics in particular. The architecture presented in this article is essentially based on IoT applications and big data analytics and therefore enables the evaluation of large amounts of data and the generation of user-defined KPIs in real-time. The digital twin is thus an essential component of these architectures, as it enables an extensive exchange of information (Mičieta, Gašo and Krajčovič, 2014 cited in Furmann, Furmannová and Więcek, 2017, p. 208).

2 Research Design

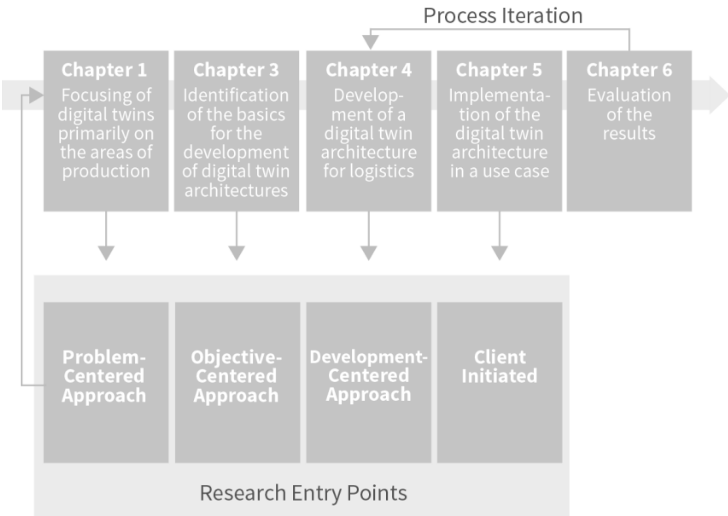

Figure 1: The DSRM for the development of a digital twin architecture in logistics based on Peffers et al. 2007

The structure of this paper and the approach to the development of a digital twin architecture are based on the Design Science Research Methodology (DSRM) for Information Systems Research according to Peffers et al. (2007). This methodology is structured into six different steps and begins with a problem-centered approach that identifies and motivates a problem. In the second step, objectives of the solution are presented, in which the necessary solution approaches are determined. Based on steps one and two, a central artifact is designed and developed in step three. The fourth step is the demonstration of the artifact in a specific context. Steps five and six evaluate and finally communicate the solution. They represent evaluation and communication, and are used for process iterations, which in turn influence the structure of the central artifact (Peffers, et al. 2007, pp. 12-14). As shown in Figure 1, the sixth step of this paper is excluded because communication has not yet taken place.

3 State of the Art of Real-Time Data Processing in Logistics

Following the DSRM by Peffers et al., the problem identification presented in chapter 1, according to which digital twins are mainly found in the area of manufacturing, is the motivation for considering this technology in the context of logistics. It is also about the question whether current architectures are suitable for an application in the field of logistics.

3.1 Internet of Things and Big Data in Logistics

Digitalization and the associated digital transformation of processes affect almost all areas of the economy and industry (Kersten, et al., 2017, p. 8). For its implementation new technological concepts are needed, which primarily relate to data management and analytics. These include comprehensive sensor technology, which serves as a data source for monitoring and improvement, as well as predictive analyses and artificial intelligence, which form the basis for the optimization of logistics processes (Kersten, et al., 2017, p. 12). The core technologies are therefore the realization of extensive sensor technologies and the development of algorithms capable of processing large amounts of data. IoT and Big Data technologies are proving to be the most promising way to process large amounts of data in real-time (Malek, et al., 2017, p. 429). The ability to extract already processed raw data in the format of KPIs from running processes and to visualize them in real-time will bring fundamental improvements in the area of data management. This is especially true in the field of logistics, where many data sets are generated (Wang, et al, 2016, p. 104).

IoT applications and Big Data analysis in combination form a considerable potential for various applications in the field of data management. In general, IoT refers to the vision of a continuous networking of objects so that these objects can communicate with each other and with their environment (Bousonville 2017, p. 25). In this context, IoT refers to a network of sensors by which data can be obtained from various processes (Hopkins and Hawking 2018, p. 576-578). IoT applications thus form the basis for comprehensive data generation. The number of data collected is substantial, depending on the area of application. However, data collection is only the first step.

At this point, it is still completely unclear what value the generated data has for the processes from which it was collected. This means that data from different sources must be merged in order to be processed further, which requires Big Data analysis (Bousonville 2017, p. 25). The term Big Data can essentially be described with the 4 Vs that stand for Volume, Velocity, Variety, Value (Dijcks 2014, pp. 3-4). Value is a particularly im-portant parameter in this context, since the analysis of large amounts of data must focus on the aspect of generating only data of relevance (Dijcks 2014, p. 4; Bousonville 2017, p. 26). The combination of IoT applications and big data analytics is done with architectures that enable end-to-end data management from data collection through data preparation up to data visualization.

Such architectures provide components for three stages (Malek, et al., 2017, p. 431): Data acquisition, data processing and data visualization.

This becomes more obvious with the consideration of an exemplary IoT architecture in Figure 2 (ISO/IEC 2016, p. 41). It shows the Inside Domain Functions of an IoT architecture, which was developed for the definition of an IoT Reference. The lowest level refers to the Physical Entity Domain, which

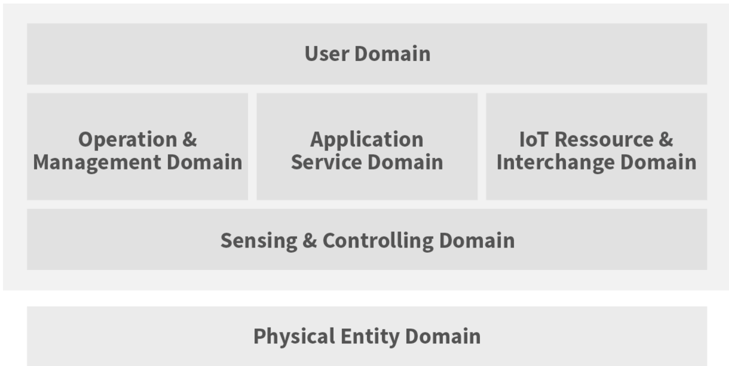

Figure 2: Exemplary IoT architecture according to (ISO/IEC 2016, p. 41)

describes the physical asset under consideration. The next level covers the Sensing and Controlling Domain, which in relation to Malek et al. represents the level of data acquisition. The data processing level is divided into the three domains Operation and Management, Application Service, IoT Resource and Interchange. This level contains the essential analysis functions for real-time data processing. The last level shows the user domain, which enables data visualization. However, a more detailed look reveals that the basics for an exact semantic description of systems is missing, as it is specifically necessary for the area of logistics.

3.2 Digital Twins in Logistics

The continuous development in the field of big data analytics and sensor technologies and the associated progress in the field of IoT technologies finally led to the advancement in digital twins (Tao, et al., 2019, p. 2405). With these platform architectures, the collected data can be processed and visualized in real-time. They provide a technical framework on which concrete industrial applications can be developed. Nevertheless, it is evident that concrete use cases for the implementation of such architectures are rarely found in logistics, although the implementation of IoT technologies in logistics offers considerable additional value (Hopkins and Hawking 2017, p. 579). If logistics objects are equipped with a comprehensive sensor system, a digital image of the respective logistics objects is created, a so-called digital twin. More precisely, a digital twin is the digital representation of a physical asset (Wohlfeld 2019, p. 65). Digital twins enable the connection between the physical and digital world, which must be based on a complete database (Tao, et al., 2019, p. 2405; Wohlfeld 2019, p. 65).

A digital twin is far more than the digital representation of a physical asset. A digital twin represents a comprehensive physical and functional depiction of an asset that provides all the information necessary to process it throughout its lifecycle. (Boschert and Rosen, 2016, p. 59). The exact definition of a digital twin depends on the integration level. A distinction must be made between a digital model, a digital shadow and a digital twin. The core of the consideration is in all cases a physical and a digital object. In a digital model, there is only a manual data flow between the physical and the digital object. A change in the physical object has no effect on the digital object and vice versa. In a digital shadow, there is an automatic data flow in at least one direction, whereby the change of the physical object leads to a change of the digital object. However, this does not apply to the reverse case. In a digital twin, the data flow between the two objects is automatic. Thus, a change to the physical object leads directly to a change to the digital object and vice versa. (Kitzinger et al. 2018, p. 1017).

The use of digital twins enables real-time communication between assets and different systems. With regard to logistics, data collection alone does not represent a major challenge. The decisive factor here is how this data must be further processed in order to offer real added value. In this context, the added value is created with the help of KPIs tailored precisely to the respective application. Depending on the sensors used, different KPIs can be determined from the same data sources in real-time, exactly as required for the respective process. This technology thus offers consider-able potential for logistics and contributes to targeted decision-making (Wang et al. 2016, p. 99).

3.3 Applicability of Lambda Architecture in Logistics Systems

The Industrial Internet of Thing (IIOT) produces massive quantities of sensor data, which arrives in a streaming fashion. The lambda architecture is an efficient big data solution for generic, scalable and fault-tolerant data processing (Gröger, C. 2018).

In the context of IoT data processing, two layers of the lambda architecture consume incoming data simultaneously. The batch layer enables time-consuming analyses on stored raw data, therefore the results are provided to the serving layer and can be consumed by the users. Using a distributed storage topology, the vast amount of sensor data is stored in the batch layer efficiently. In the meanwhile, the speed layer enables real-time analysis of the incoming data streams.

Because of the limitation of computational resources, it is often impossible to load whole datasets at once and analyzing them with classical machine learning models. On the contrary, visiting each instance of the data stream only once and analyzing it with either an adaptive online model or a robust batch model provides the chance to get fresh knowledge from data streams in time, which is of vital importance for IoT data analysis applications.

3.4 Research Gap on Digital Twin Architectures in Logistics

As already pointed out at the beginning of this article, the majority of publications on digital twins relate to the area of shop floor and production management (Zhuong, et al. 2018, p. 1153; Qi, et al. 2018, p. 238). Kitzinger et al. (2018) offer a comprehensive literature analysis on this topic. It shows that more than half of the publications on the subject of digital twin initially describe basic concepts only. Just a quarter of the publications refer to concrete use cases, but most of them are in the area of simulation, product lifecycle management and manufacturing in general (Kitzinger et al. 2018, pp. 1018-1020).

In the context of logistics, Hopkins and Hawking contribute to the application of IoT and Big Data Analytics in logistics. This contribution is based on a literature review on these topics. The result of this investigation is a lack of concrete use cases of both topics, from which the claim is derived to close the gap between theory and logistics practice. Finally, a Big Data Framework is examined using a case study approach. However, this study does not explicitly focus on digital twins, but on the influence of IoT and Big Data on various problems in transport logistics (Hopkins and Hawking 2018). By looking at existing reference architectures, it also becomes apparent that these are too generic to be used in logistics. Furthermore, a functionality to consider a semantic description of the systems to be considered is missing.

Thus, research in the field of digital twins is at an early state, where the main objective is to find new applications for this technology (Negri, et al. 2017 p. 946). Therefore, this paper makes a contribution to the use of digital twin architectures in logistics. It shows how a digital twin architecture can be set up to achieve a seamless integration of logistics systems.

4 A Digital Twin Architecture for Logistics

Considering Big Data Analytics in logistics as well as existing architectures, a digital twin architecture for logistics is now being developed in the third step of the DSRM according to Peffers et al. Therefore, this architecture forms the central artifact. Figure 3 shows a real-time IoT data processing and analyzing platform with a lambda architecture, which aims to provide a scalable and powerful infrastructure for IoT data acquisition, processing and visualization. As an IoT solution for logistics, it is flexible and industrial-application-oriented. The architecture is composed of four layers, as described in detail in the next sections. It is a digital twin architecture with an optional data acquisition layer. The digital twin architecture itself has layers for data visualization, data processing and a semantic layer for providing the overall system model and data integration. These layers are used to enrich, integrate and process the data from the sensors to values that are finally visualized in real-time.

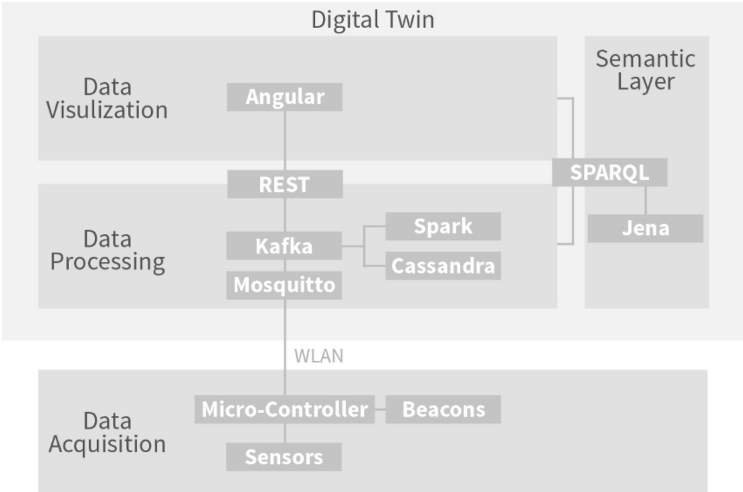

Figure 3: Lambda architecture for real-time IoT analytics in logistics

The implementation of this lambda architecture was realized with a modified SMACK (Spark, Mesos, Akka, Cassandra and Kafka) stack, which is a proven distributed big data stack for streaming data processing. The following sections describe how the individual layers of this infrastructure operate. Furthermore, it is shown which software tools were used to implement these layers and why these software components are best suited for this layer.

4.1 Data Acquisition

With regard to data acquisition, the architecture described here is sensor-independent. Therefore, the description of data collection in this article is a secondary aspect. The selection of a sensor system ultimately decides on the limitation of the possible analyses, since these can only be as good as the sensor system itself. For data acquisition it is useful to select a microcontroller, which can to hold a multitude of different sensors. A decisive aspect in the selection of these modules is the costs. It is important to emphasize that even with a simple sensor system the most diverse evaluations are possible. When such microcontrollers and sensors are used, it must always be ensured that they are also suitable for industrial applications. They must be resistant to vibrations and temperature fluctuations.

4.2 Data Processing

The architecture shown in Figure 3 uses a modified SMACK stack to perform real-time and batch KPI analysis such as shock detection, indoor localization, and usage analysis. Unlike the classic SMACK stack, Apache Akka has been replaced by Apache Nifi, which provides similar features with a more straightforward structure. In addition, several backend functions have been implemented with the Java Spring Boot Framework.

The sensor data is transferred to the infrastructure via the microcontroller. These raw data arrive in the MQTT data broker Mosquitto. There the raw data is pre-processed and distributed to different target units. Kafka transforms the data streams in the overall system and thus forms a distributed data processing platform that enables real-time data stream processing with high throughput. Streaming data is also stored in a Cassandra database, an extensive NoSQL database, for batch analysis. Spark is used as a real-time data analysis engine where the data stream is analyzed in near real-time using the native MLlib machine learning library. After backend processing, the raw data stream and analysis results are visualized on the web frontend.

Data processing is used in industrial applications to generate the KPIs required for the respective process in real-time. The definition, calculation and visualization of KPIs for a specific application is, therefore, the central analysis function of a digital twin system. The combination of lambda architecture and digital twin enables powerful and scalable KPI calculations in real-time. The KPIs generated by this kind of architecture enables companies to quickly determine the condition of their assets. Three steps are required to define and store a new KPI function for a specific scenario:

1. Implementation of the KPI function
2. Implementation of KPI visualization
3. Adding a semantic description to an ontology

The KPI functions are calculated with statistical and machine learning models in batch or real-time. A distinction must be made between whether it is really necessary to generate a KPI in real-time, or not. In general, each KPI function is visualized on the frontend, allowing the user to monitor all relevant indicators. Thus, it makes sense to build the frontend of such an architecture component-based. This means that each KPI function is organized as an isolated component, which makes it much easier to implement new KPIs into the architecture. The components communicate with the backend via a REST API. The real-time KPIs are visualized dynamically from the streaming data. After the implementation of a new KPI function and the visualization of this indicator, the relevant physical objects as well as the analysis functions are described in the ontology within the semantic layer. The concrete advantage here is the standard semantic annotation in an overall model.

4.3 Data Visualization

In addition to data processing, data visualization is another important component, since end users have access to the processed data here. KPIs and the digital description of physical objects are visualized on the frontend. Finally, the optimization options can be identified on the basis of the data visualized in the frontend. For example, Angular is used to create a component-based Web interface. This also enables flexible extensibility of the functions for the frontend. The raw data and analysis results are transferred to the frontend in the data stream and displayed dynamically so that the user can monitor the systems according to the real-time conditions. In this context, it is useful to display the key figures and graphical evaluations on a mobile device, since these are particularly suitable for monitoring running processes. A large number of different KPIs that are relevant for an application in the logistics context can be displayed on the frontend or on the user interface.

4.4 Semantic Layer for Digital Twins in Logistics

The introduction of digital twins faces difficulties due to a lack of semantic interoperability between architectures, standards and ontologies (Datta 2016, p. 1). A digital twin needs a detailed model of its physical counterpart and its relevant environment. This can be a business-oriented semantic model to provide an integrated view of all relevant units in detail, based on the use and extension of standard ontologies. This includes, for example, the relevant assets of the company for which a digital twin is defined and the microcontrollers and sensors associated with those assets. For example, the heterogeneity of the various sensors used is managed with a standard sensor ontology such as W3C SSN (Haller, et al., 2017). New sensors and new assets can be easily connected and configured by instantiating ontology concepts.

The semantic layer of a digital architecture mainly consist of software components and ontologies in a RDF format. The software components are primarily Triple Store and Reasoner. These ontologies are stored in the Triple Store and are used by semantic SPARQL queries executed by the Reasoner. To keep license costs low, open source software can be used, such as the free open source Jena Framework (Apache Software Foundation 2019), which can serve as the basis for the implementation of a se-mantic layer. The composed ontology is a structured semantic model of all relevant entities such as IoT devices, assets and their relationships. The top level of the ontology architecture describes the digital twin and its analyses. The company and its assets follow in the next ontology layer. To support the digital twin with values, IoT devices connected to assets are described in the lowest ontology layer.

4.5 Research Progress by the Presented Architecture

The concept of a digital twin architecture is a very flexible and cost-effective IoT solution. In order to become flexible, analytical modules with logic and display functionality are semantically combined in a lambda architecture. For each KPI function on the frontend, there is a corresponding semantic description in the semantic layer. In the Triple Store, not only the digital description of physical objects is stored, but also respective services of the object, for example the machine learning model used for this object and the relevant sensor types. Furthermore, the semantic model can be easily updated when changes are made to sensors or when machine learning models are updated.

Another aspect is the use of modular micro services. The analysis functions of digital twin architecture can be implemented as modular micro-services, with semantic annotation. The analysis functions and frontend services are fully customizable. For changing sensors or analysis models, the micro-services can be easily extended by adapting semantic annotations and function changes. Also a component-based user interface with Angular has been implemented for flexible customization. For installation, the entire application is packaged in docker images so that it can be used on different platforms at any time.

The large data tools used for the architecture are flexibly scalable. As a large data analysis solution for industrial applications, a digital twin architecture is suitable for various scenarios and applications. Depending on the type and number of sensors, the complexity of the analysis models and the availability of computing resources, the performance of a digital twin architecture can be adapted to the respective application area. By using a combination of these distributed large data tools, a certain fault tolerance is ensured by storing computing information redundantly over different computing nodes. This ensures that the data is not lost in the event of a system failure. Communication between the components is usually implemented either with the REST API or with specific connectors, which are also easily expandable.

The flexibility of this architecture is further demonstrated by its independence from the sensor technology used. The specific sensor types can vary from user to user. The metadata of sensors and analysis functions are stored and linked in the semantic layer. The Data Broker receives all data and its metadata in JSON format, so that further processing and analysis is planned on the basis of metadata and information from the semantic layer.

5 Application Scenario of the Digital Twin Architecture for Real-Time Data Processing Based on Artificial Intelligence

In order to demonstrate the framework presented in section 4, RIOTANA was developed in the context of logistics, which contains all described properties of the digital twin architecture presented here. According to the DSRM by Peffers et al., in chapter 5 the implementation and demonstration of the artifact, the digital twin architecture developed here, is realized. RIOTANA stands for real-time Internet of Things analytics and represents a digital twin architecture with which KPIs can be updated in real-time.

In intralogistics it is possible to transform industrial trucks into a virtual asset with the help of comparatively inexpensive sensor technology. Hence, analyses can be carried out with which related internal processes can be fundamentally optimized. With RIOTANA, a comprehensive forklift control system can be implemented without having to access the elec-tronics of the industrial trucks. Based on discussions with stakeholders and the analysis of existing forklift control systems, a comprehensive system for controlling forklift fleets was developed. The modular structure of the architecture makes changes easy to implement, as the mathematical calculations implemented in the architecture can be adapted to new application conditions and new sensors. The sensor modules collect data on position, acceleration and localization, for example. These different data types are merged by the RIOTANA architecture, processed in real-time and displayed on a web frontend.

Already with these three sensor types in combination, various analyses are possible, which offer a complete overview of the running processes. This includes the current workload, the current location as well as detected shocks and collisions. Accordingly, an entire forklift fleet can be equipped with such sensors, which in turn allows conclusions to be drawn about the overall effectiveness of this fleet.

Figure 4 shows a section of the RIOTANA user interface during a field test. This field test was particularly concerned with a test of the implemented machine learning algorithms for the detection of shocks, as this is an important KPI for a large number of industrial scenarios that must be generated in real-time. Especially this indicator is regarded as a classification problem, for which a K-Nearest Neighborhoods (KNNs) model is used. This is done by collecting time series data from motion sensors attached to the devices. For this purpose, a sliding window is transferred to the time series, whereby each window is regarded as a pattern. Shock patterns are classified on the basis of standard patterns, allowing each incoming pattern to be classified in real-time.

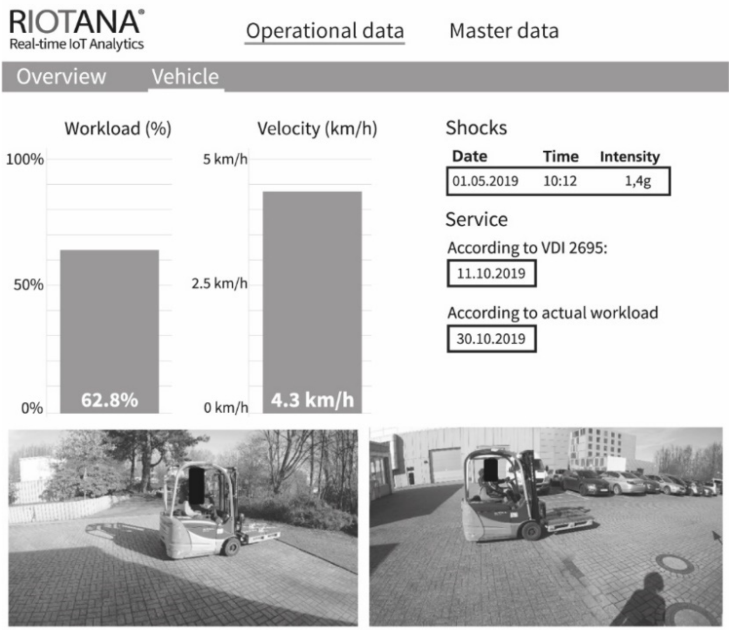

Figure 4: The RIOTANA user interface during a field test

Using the sensor modules RIOTANA can be integrated into existing processes as a "retrofitting solution". By using cheap sensor technology, previously "non-intelligent" objects can become virtual assets that are available for a variety of analyses and can be used to optimize processes. In addition, the use of the sensor modules is optional. RIOTANA can also be integrated into processes in which data are already collected but not yet evaluated. The application of such architectures is conceivable on the basis of digital platforms that operate according to the "as-a-service" principle (Otto, et al. 2019, p. 115).

6 Conclusion and Future Work

In this paper a digital twin architecture was presented, which enables the analysis and processing of large amounts of data in real-time on the basis of IoT applications and big data analytics. It was also shown that the realization of such architectures can be realized with open source software components (Holtkamp 2019, p. 10). The special feature is the description of a digital twin architecture with reference to a concrete application in logistics. It is exactly this practical relevance that presents a particular challenge in the further development of this architecture. This is expressed in an iterative process according to DSRM by Peffers et al., shown in Figure 1. There have to be further investigations on how such architectures can be used in logistics, which in turn has an influence on the structure of the architecture presented here.

The collection of data in an industrial context is always a critical topic that must be considered with special attention. This is particularly the case for personal data. In order to ensure that the processed data is only made available to those who are authorized to do so, a corresponding sensor connector must be implemented in the sensor module. In this way, access to the data can be considerably restricted.

Another important technical aspect is the further development of the machine learning functions in RIOTANA in order to achieve even more precise results with the shock detection. In addition to the further development of machine learning functions to recognize patterns and anomalies and the implementation of software components to ensure data sovereignty, there are also conceptual questions. These include questions about the criteria that determine whether an asset needs a digital representation. Furthermore, it will be necessary to clarify which processes or systems require real-time data processing at all. Beyond that, there are no descriptions of how such architectures can be implemented in processes. Finally, it becomes evident that due to the focus of digital twins on the area of manufacturing, further investigations are necessary with regard to logistics.

References

Apache Software Foundation, 2019. Apache Jena - A free and open source Java framework for building Semantic Web and Linked Data applications. Available at: <https://jena.apache.org/index.html> [Accessed 3 April 2019].

Boschert, S., Rosen, R. 2016. Digital Twin-The Simulation Aspect. In: Hehenberger, P., Bradley, D.: Mechatronic Futures - Challenges and Solutions for Mechatronic Systems and their Designers. Springer, Int. Pub. Switzerland, 2016, pp. 59-74.

Bousonville, T., 2017. Logistik 4.0. Die digitale Transformation der Wertschöp-fungskette. Wiesbaden: Springer Gabler.

Chae, B., 2009. Developing key performance indicators for supply chain: an industry perspective. In: Supply Chain Management: An International Journal, [e-journal] 14 (6), pp.422 - 428. https://doi.org/10.1108/13598540910995192

Datta, S.P.A., 2016. Emergence of Digital Twins. MIT Auto-ID Labs, Massachusetts Institute of Technology, Available at: <https://dspace.mit.edu/bitstream/handle/1721.1/104429/Digital%20Twins.pdf?sequence=152> [Accessed 3 April 2019].

Dijcks, J.-P., 2014. Oracle: Big Data for the Enterprise. [pdf] An Oracle White Paper from September 2014. Available at: <https://www.oracle.com/assets/wp-bigdatawithoracle-1453236.pdf> [Accessed 3 Mai 2019].

Furmann, R. Furmannová, B., Więcek, D., 2017. Interactive design of reconfigurable logistics systems. In: Procedia Engineering, [e-journal] 192, pp. 2017 - 212. https://doi.org/10.1016/j.proeng.2017.06.036

Gröger, C., 2018. Building an Industry 4.0 Analytics Platform. In: Datenbank Spektrum, [e-journal] 18 (1), pp. 5 - 14. https://doi.org/10.1007/s13222-018-0273-1.

Gupta, A, Thakur, H. K., Shrivastava, R., Nag, S., Kumar, P., 2017. A Big Data Analysis Framework Using Apache Spark and Deep Learning. In: IEEE, International Conference on Data Mining Workshops (ICDMW). New Orleans, LA, USA, 18-21 Nov 2017. https://doi.org/10.1109/ICDMW.2017.9.

Haller, A.; Janowicz, K.; Cox, S.; Le Phuoc, D.; Taylor, K., Lefrançois, M., 2017. Semantic Sensor Network Ontology. W3C Recommendation. Available at: <https://www.w3.org/TR/vocab-ssn/> [Accessed 3 April 2019].

Holtkamp, B., 2019. RIOTANA: Echtzeit-Ermittlung von Kennzahlen in Produktion und Logistik. Ein leichtgewichtiger Einstieg in Industrie 4.0 für den Mittelstand. [pdf] Fraunhofer Institute for Software and Systems Engineering: Logistics. Available at: < https://www.isst.fraunhofer.de/de/geschaeftsfelder/digitization-in-logistics/RIOTANA.html> [Accessed 3 Mai 2019].

Hopkins, J., Hawking, P., 2018. Big Data Analytics and IoT in logistics: a case study. The International Journal of Logistics Management, [e-journal] 29(2), pp.575 - 591. https://doi.org/10.1108/IJLM-05-2017-0109.

ISO/IEC, 2016. ISO/IEC CD 30141:20160910(E) [pdf]. Available at: <https://www.w3.org/WoT/IG/wiki/images/9/9a/10N0536_CD_text_of_ISO_IEC_30141.pdf> [Accessed 16 May 2019].

Jeske, M., Grüner, M., Weiß,F., 2014. White Paper: Big Data in der Logistik. [pdf] T-Systems. Available at: <https://www.t-systems.com/blob/219384/067ce0047c88fc357ee42b102204512d/DL_WP_Big_Data_in_Logistik.pdf> [Accessed 3 Mai 2019].

Hazen, B.T., Skipper, J.B., Ezell, J.D., Boone, C.A., 2016. Big data and predictive analytics for supply chain sustainability: A theory-driven research agenda. Computers & Industrial Engineering, [e-journal] 101, pp. 592 - 598. http://dx.doi.org/10.1016/j.cie.2016.06.030.

Kersten, W., Seiter, M., von See, B., Hackius, N., Maurer, T., 2017. Trends und Strategien in Logistik und Supply Chain Management. Chancen der digitalen Transformation. Bremen: DVV Media Group GmbH.

Kitzinger, W., Karner, M., Traar, G., Henjes, J., Sihn, W., 2018. Digital Twin in manufacturing: A categorical literature review and classification. In: IFAC-PapersOnLine, [e-journal] 51 (11), pp. 1016 - 1022. https://doi.org/10.1016/j.ifacol.2018.08.474.

Malek, Y.N., Kharbouch, A., El Khoukhi, H., Bakhouya, M., De Florio, V., El ouadghiri, D., Latre, S., Blondia, C., 2017. On the use of IoT and Big Data Technologies for Real-time Monitoring and Data Processing. In: Procedia Computer Science, [e-journal] 133, pp.429 - 434. https://doi.org/10.1016/j.procs.2017.08.281.

Mishra, N., Lin, C.-C., Chang, H.-T., 2014. A Cognitive Oriented Framework for IoT Big-data Management Prospective. http://dx.doi.org/ 10.1109/IC-CPS.2014.7062233.

Negri, E., Fumagalli, L., Macchi, M., 2017. A Review of the Roles of Digital Twin in CPS-based Production Systems. A Review of the Roles of Digital Twin in CPS-based Production Systems. In: Procedia Manufacturing, [e-journal] 11, pp. 939 - 948. https://doi.org/10.1016/j.promfg.2017.07.198.

Otto, B., ten Hompel, M., Wrobel, S., 2019. Industrial Data Space. Referenzarchitektur für die Digitalisierung der Wirtschaft. In: R. Neugebauer, ed. 2019. Digitalisierung. Schlüsseltechnologien für Wirtschaft & Gesellschaft. München: Springer Vieweg. pp. 113-133.

Park, H., Srinivasan, A., Bolsius. R., Hsiao, E., Toillion, J. Shukla, V., Oracle International Corporation 2014. Real-Time Monitoring Event Analysis and Monitoring. US 2014/0095537 A1 (Patent Application Publication).

Peffers, K; Tuunanen, T., Rothenberger, M. A., Chatterjee, S., 2007. A Design Science Research Methodology for Information Systems Research, [e-journal] 24 (3), pp. 45 - 77. https://doi.org/10.2753/MIS0742-1222240302.

Qi, Q., Tao, F., Zuo, Y., Zhao, D., 2018. Digital Twin Service towards Smart Manufacturing. In: Procedia CIRP, [e-journal] 72, pp. 237 - 242. https://doi.org/10.1016/j.procir.2018.03.103.

Spangenberg, N., Roth, M., Mutke, S., Franczyk, B., 2017. Big Data in der Logistik. Ein ganzheitlicher Ansatz für die datengetriebene Logistikplanung, -überwachung und -steuerung. Industrie 4.0 Management, [e-journal] 33(4), pp. 43-47.

Tao, F., Zhang, H., Liu, A., Nee, A. Y.C., 2019. Digital Twin in Industry: State-of-the-Art. IEEE Transactions on Industrial Informatics, [e-journal] 15(4), pp.2405-2415. https://doi.org/10.1109/TII.2018.2873186

Wang, G., Gunasekaran, A., Ngai, E. W. T., Papadopoulos, T., 2016. Big data analystics in logistics and supply chain management: Certain investigations for research and applications. Int. J. Production Economics, [e-journal] 176, pp.98-110. https://doi.org/10.1016/j.ijpe.2016.03.014

Wohlfeld, D., 2019. Digitaler Zwilling für die Produktion vom Übermorgen. Große Fortschritte auf dem Forschungscampus ARENA2036. Zeitschrift für wissenschaftlichen Fabrikbetrieb, [e-journal] 144 (1-2), pp.65-67. https://doi.org/10.3139/104.112008.

Zhuang, C., Liu, J., Xiong; H., 2018. Digital twin-based smart production management and control framework for the complex product assembly shop-floor. In: The International Journal of Advanced Manufacturing Technology, [e-journal] 96, pp. 1149–1163 https://doi.org/10.1007/s00170-018-1617-6.