IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.Doi Number

Digital Immune Systems for Civil Infrastructure

Kay Smarsly
Institute of Digital and Autonomous Construction, Hamburg University of Technology, Hamburg, Germany
United Nations University (UNU) Hub on Engineering to Face Climate Change, United Nations University Institute for Water, Environment and Health (UNU-INWEH), Hamburg University of Technology, Germany

Corresponding author: Kay Smarsly (e-mail: kay.smarsly@tuhh.de).

This work was supported by the German Research Foundation (DFG) under grant SM 281/44-1.

ABSTRACT Structural health monitoring (SHM) has become a central tool for managing civil infrastructure. In modern SHM, the physical structure and the SHM system together form a coupled cyber-physical system. Due to the networked nature of the cyber-physical system, threat categories no longer only comprise anomalies in the physical structure or internal SHM system faults, but also external cyberattacks. The combination of all capabilities required to withstand such threats can be subsumed under the term "digital immune system". This paper proposes a biologically inspired, computing-oriented reference architecture for digital immune systems for civil infrastructure. The reference architecture is derived from an analysis of the functions and components of the biological immune system, following a design science research methodology grounded in structured functional abstraction. The paper presents a descriptive, architecture-level formal specification of the reference architecture and threat categorization, thereby providing a rigorous conceptual foundation for future modeling and simulation of digital immune systems in cyber-physical SHM. The formalization defines structural constraints of the reference architecture without imposing implementation-specific correctness criteria. To illustrate operational plausibility without prescribing a concrete implementation, the paper additionally presents an illustrative use-case scenario based on an existing SHM test setup. The paper concludes with recommendations for implementing and validating digital immune systems in real SHM deployments. Overall, digital immune systems have the potential to enhance the resilience of cyber-physical SHM systems for civil infrastructure and to stimulate interdisciplinary research at the interface of civil engineering, computer science, and immunology.

INDEX TERMS Structural health monitoring, cyber-physical systems, artificial immune systems, civil engineering, fault diagnosis.

I. INTRODUCTION

Structural health monitoring (SHM) is an essential instrument for ensuring the safety and functionality of civil infrastructure, which is increasingly subjected to aging as well as increasing traffic and environmental impacts [1]. Advances in sensing technology, wireless communication, and embedded computation have enabled continuous or near real-time monitoring of structural response under real operating conditions by means of wireless sensor networks deployed on civil infrastructure, which generate large-scale data streams that are analyzed using computational modeling and simulation pipelines [2]. As SHM systems have become tightly integrated with communication networks and cloud services, SHM systems have inherited the cybersecurity vulnerabilities of networked cyber-physical systems, such as exposure to attacks on communication links, embedded devices, and cloud backends [3], for example with the aim of conducting sabotage or data theft. Studies of smart structures have highlighted that compromised monitoring or control channels may lead to unsafe operating states particularly in critical infrastructure, e.g. if attackers tamper with sensor data or actuator commands on bridges or on wind turbines subjected to wind-induced loads [4]. Recent work has shown that data-driven (e.g., machine-learning-based) SHM models can be manipulated by adversarial modifications of sensor data, and that such perturbations can induce misclassification of damaged and undamaged states [5]. Related findings from adjacent critical-infrastructure domains, particularly IoT-based smart grids, indicate that cyberattacks can remain stealthy even when attackers possess only limited system knowledge. For instance, zero-parameter-information data-integrity attacks illustrate how state estimation may be manipulated without requiring complete knowledge of the target system [6]. At the same time, recent review studies

indicate that machine-learning-based operational analytics are susceptible to adversarial manipulation [7]. Additional research has introduced secure state-estimation approaches based on hybrid homomorphic encryption [8]. Furthermore, coordinated cyber-physical attacks have been shown to combine the manipulation of control actions with the falsification of measurements in a stealthy and effective manner [9].

Overall, in the context of SHM, the "health" no longer denotes just the ability of a structure to fulfill its intended function safely and reliably [10], but also includes the health of the SHM system itself. In other words, the physical structure and the SHM system form a coupled cyber-physical system that must be considered and protected as a whole. Next-generation SHM therefore requires additional protective layers, such as secure communication and authentication mechanisms, redundancy and validation in sensing and data processing, and adaptive anomaly detection that maintains credible estimates of structural health even under cyberattacks or sensor faults. The aforementioned capabilities are often summarized under the term "digital immune systems". Herein, the term "digital immune system" refers to the digital, data-driven layer of this coupled system, i.e. to sensing, data processing, and system-level response functions, rather than to direct physical intervention on the monitored structure.

Digital immune systems (DIS), sometimes referred to as "artificial immune systems" (AIS), have been explored for decades outside SHM as biologically inspired architectures for adaptive detection, protection, and decision-making across multiple domains. In the banking domain, for example, an artificial immune system is proposed for identifying fraudulent credit card transactions [11]. In networked computing environments, DIS-based concepts are used for providing adaptive, self-monitoring protection mechanisms [12]. In IT security, DIS typically denotes large-scale, artificial-immune-system-based defense architectures for complex information networks [13]. In industrial fault detection and condition monitoring, artificial immune systems are applied to fault detection and isolation in machinery and process plants and are commonly classified as DIS [14]. In immunogenomics, the term refers to high-resolution in silico models of the human immune system derived from molecular and genomic data [15]. In communication network analysis and protection, immunization strategies for complex networks are explicitly framed as steps toward designing and deploying digital immune systems [16].

Among the application areas listed above, cybersecurity represents one of the most prominent and mature fields for digital immune systems, as the intrinsic capabilities for adaptive anomaly detection, self-learning defense, and distributed protection align closely with the requirements of modern cyber-physical and information systems. Survey articles within the domain of cybersecurity show that digital immune systems form an established line of work, which incorporates key immunological principles, such as clonal expansion, danger signaling and self/non-self discrimination – the rejection of detectors that match normal ("self") behavior so that only detectors for previously unseen ("non-self") patterns remain – in a strongly abstracted computational form [17]. Negative-selection methods apply the principle of self/non-self discrimination from immune systems to the task of anomaly detection in system calls and network traffic, while clonal-selection-based approaches refine detector populations by clonal expansion and mutation to represent memory effects [18]. Specific architectures include immunity-based security layers for Internet protocols, which embed immune mechanisms into protocol stacks [19], danger-theory-inspired models that focus on danger signals rather than solely on self/non-self discrimination [20], and dendritic-cell-based algorithms that operationalize aspects of innate immune processing for anomaly detection in artificial immune systems [21].

At the other end of the spectrum of biological grounding stands Gartner's use of the term "digital immune system", which describes a management and architectural framework for software landscapes that combines observability, automated and AI-assisted testing, chaos engineering, incident response automation, and integrated development, security and operation workflows rather than explicit immunological mechanisms [22]. A similarly loose use of the term DIS appears in parts of the commercial security industry, where products are marketed as "enterprise immune systems" or as providing "digital antibodies". In practice, the aforementioned products rely on machine-learning-based anomaly detection methods similar to methods developed in the technical literature [23].

In summary, the terms "digital immune system" and "artificial immune system" are used across application domains with varying degrees of biological abstraction. While existing research has developed a wide range of biologically inspired algorithms, including clonal selection, negative selection, and dendritic-cell-based approaches, these contributions predominantly address algorithmic aspects, and in the context of cyber-physical SHM of civil infrastructure, a lack of biologically inspired digital immune systems, formulated as comprehensive, architecture-level frameworks, persists.

This paper proposes a biologically inspired computational reference architecture for digital immune systems of civil infrastructure, developed following a design science research methodology [24]. Unlike classical artificial immune systems, which primarily focus on algorithmic anomaly detection mechanisms, the present work develops an architectural integration framework for cyber-physical SHM systems grounded in biological immune principles. In addition, the reference architecture and the underlying threat categorization are formally specified in many-sorted first-order logic [25]; a formalization that enables an explicit separation of entity types, prevents category inconsistencies within the architectural specification, and remains comprehensible to

civil engineers with limited formal logic background. Building upon the biological immune system as conceptual model, the proposed architecture is structured into an innate layer and an adaptive layer. The innate layer comprises modules for physical protection (e.g., coatings and sealants), local sensing, non-specific diagnostics, feature processing, and threat containment, together with global amplification logic and regulatory signals that distribute alerts throughout the system. The adaptive layer includes adaptive response modules (providing case-specific response orchestration, response regulation, threat neutralization and intervention memory) as well as global elements consisting of markers and learners that allow efficient adaptation to threats acting on the coupled system of physical structure and SHM system. All elements communicate via communication links (equivalent to nerves in the biological immune system). The coordination can be managed by a central server (equivalent to the central nervous system), while execution may be realized in a distributed manner on either fixed or mobile sensor nodes (i.e., mobile robots). The formalization presented in this paper is intended as a descriptive architectural specification and does not constitute a prescriptive correctness framework for specific implementations. The proposed architecture is intended as a general, implementation-independent reference framework for cyber-physical SHM rather than as a system-specific design or prototype description.

The paper is structured as follows: Section II presents the functional principles and components of the biological human immune system that are relevant for digital immune systems for civil infrastructure. Section III develops the reference architecture and, upon analyzing the threat categories and the requirements, systematically assigns the immunological principles and components to global elements and local modules of the digital counterpart. Section IV discusses implementation-oriented views of the architecture, including a process view, a functional deployment view, and an illustrative use-case scenario, before summarizing recommendations and limitations. Section V summarizes the main results and outlines future research perspectives.

II. THE BIOLOGICAL IMMUNE SYSTEM

In this paper, the biological immune system denotes the human immune system in an abstracted biological form and is understood as a network of organs, cells and molecules that protects the organism against disease. A central task of the immune system is to distinguish between "self" and "non-self" types of structures, i.e. to tolerate endogenous structures and to eliminate foreign or altered structures that are recognized as antigens. Protection is organized within the innate and the adaptive immunity:

- *Innate immunity* comprises anatomical barriers (such as skin and mucous membranes) as the first line of defense, together with rapid immunological mechanisms that operate without prior adaptation (e.g., fever); at this stage, recognition is primarily driven by pathogen-associated molecular patterns (PAMPs), sensed via pattern recognition receptors.
- *Adaptive immunity* develops specific responses against particular antigens and forms immunological memory.

In addition, a distinction is made between *humoral components*, which act in body fluids, and *cellular components*, which are mediated by immune cells. With the aim of developing a reference architecture for digital immune systems for civil infrastructure – and for the sake of comprehensibility – only the elements of the biological immune system are considered in the following subsections that later will receive a direct functional counterpart in the reference architecture for digital immune systems for civil infrastructure (Fig. 1).

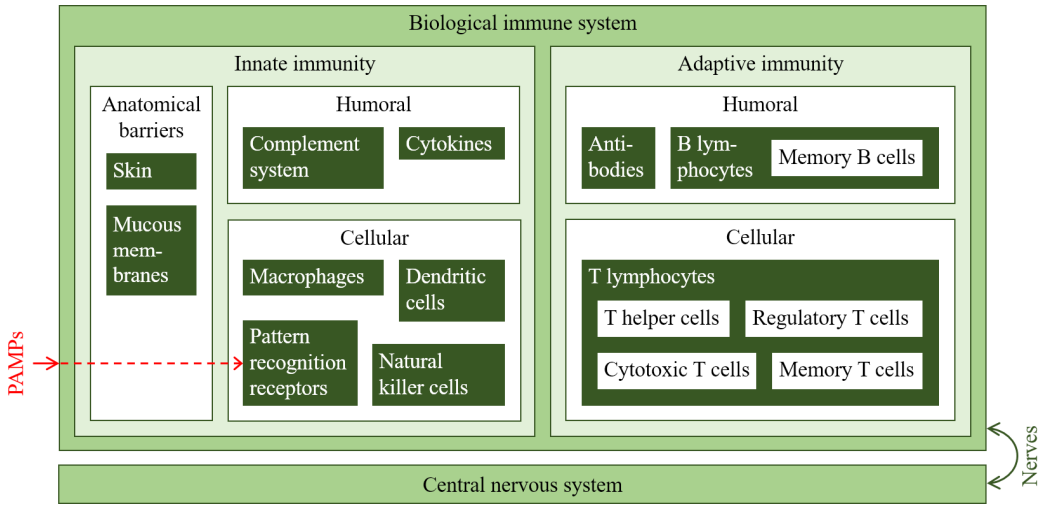

FIGURE 1. Components of the biological immune system serving as a conceptual basis for the proposed digital immune system for civil infrastructure.

A. INNATE IMMUNITY

Innate immunity denotes the part of the immune system that reacts quickly and in a largely non-specific way to pathogens and their conserved molecular patterns (e.g., PAMPs). Innate immunity is continuously active from birth and does not require prior contact with a particular infectious agent; it combines anatomical barriers and early-defense mechanisms that often control threats on their own or limit the threats until adaptive immunity becomes effective. The following subsections describe anatomical barriers as well as humoral and cellular mechanisms of innate immunity.

1) ANATOMICAL BARRIERS

Anatomical barriers form the outermost level of innate immunity. Skin and mucous membranes provide physical protection and shield the interior of the body from contact with many pathogens and pathogen-derived material (antigens). Many potentially harmful agents are removed or neutralized at this level before internal components of the immune system become involved.

2) HUMORAL INNATE IMMUNITY

Humoral mechanisms of innate immunity act in blood and tissue fluids and contribute to early reactions against pathogens by circulating through the body and reinforcing local recognition events. Humoral innate immunity includes the complement system as an effector mechanism and cytokines as soluble regulatory mediators of innate immune responses.

- The *complement system* consists of plasma proteins that circulate in body fluids in an inactive form and can be activated by conserved PAMPs on pathogen surfaces. Complement activation supports the immune response by marking antigens, promoting recruitment of immune cells and contributing directly to damage of cells or particles that carry antigens.
- *Cytokines* are soluble signaling molecules released by immune and non-immune cells after recognition of pathogens or antigenic patterns (e.g., PAMPs). Although cytokines do not act as effector molecules (such as complement), cytokines are classified in this paper as humoral components because of the action in body fluids and regulation of innate immune reactions at local and systemic levels.

3) CELLULAR INNATE IMMUNITY

Cellular mechanisms of innate immunity rely on a limited set of cell types that detect conserved molecular patterns via pattern recognition receptors and react locally to recognized structures. *Pattern recognition receptors* are expressed by immune and non-immune cells and recognize characteristic molecular patterns on pathogens or stressed/damaged cells rather than individual antigens, thereby providing the initial signal that an encountered structure falls into a "non-self" type and may represent a potential threat. The focus here lies on representative cellular components of innate immunity, i.e. macrophages, dendritic cells, and natural killer cells.

- *Macrophages* use pattern recognition receptors to detect damaged tissue and invading-material-bearing PAMPs, subsequently taking up and degrading the material, while releasing regulatory signals, such as cytokines. Macrophages therefore combine local tissue assessment, removal of suspect material, and initiation of broader responses.
- *Dendritic cells* serve as a functional bridge between innate and adaptive immunity. Upon receptor engagement, dendritic cells capture pathogens and their products, process the material into antigenic fragments, extract characteristic features, and present antigen fragments to cells of adaptive immunity. As antigen-presenting cells, dendritic cells connect innate pattern recognition to the initiation of antigen-specific adaptive responses.
- *Natural killer cells* recognize aberrant expression patterns on the surface of body cells, e.g. loss of normal markers or expression of stress-induced molecules in virus-infected cells. Natural killer cells can destroy such cells without prior antigen-specific sensitization and therefore provide a rapid cellular reaction against infected or transformed cells.

Innate immunity reacts within minutes to hours after exposure to pathogens or their molecular patterns and often limits or clears threats before adaptive mechanisms become fully engaged. Furthermore, the innate immune system stimulates the adaptive immune system. In many situations, adaptive immunity refines and stabilizes a response that innate mechanisms have already started.

B. ADAPTIVE IMMUNITY

Unlike innate immunity, adaptive immunity requires stimulation through prior exposure to infectious agents and responds with high specificity to microbial or non-microbial molecules (mostly proteins). Adaptive immunity builds on signals received from innate immunity and provides antigen-specific responses and immunological memory. Adaptive mechanisms specialize in particular antigens instead of reacting in the same way to all threats. The following subsections distinguish humoral adaptive immunity and cellular adaptive immunity.

1) HUMORAL ADAPTIVE IMMUNITY

Humoral adaptive immunity targets antigens in body fluids through antibodies. *Antibodies* (also called "immunoglobulins") are produced by differentiated B lymphocytes and bind specifically to the corresponding antigen. Antibodies neutralize antigens, mark antigens for uptake by other cells, and cooperate with the complement system. Furthermore, successful immune responses are stored in memory B cells.

- *B lymphocytes* carry antigen-specific receptors (membrane-bound antibodies) and recognize antigens together with additional activation signals. After

sufficient stimulation, B lymphocytes proliferate and differentiate to antibody-producing cells.

- *Memory B cells* arise from a fraction of activated, differentiated B lymphocytes. Memory B cells persist after clearance of the initial antigen and enable faster and stronger antibody responses upon renewed exposure to the same antigen. Due to continuous exposure to antigens, the proportion of memory B cells increases with age.

2) CELLULAR ADAPTIVE IMMUNITY

Cellular adaptive immunity uses *T lymphocytes* and several subsets thereof to recognize and eliminate antigen-bearing cells and to regulate immune responses. The following description concentrates on T lymphocytes and on distinct subsets, i.e. T helper cells, cytotoxic T cells, regulatory T cells and memory T cells. T lymphocytes recognize antigens that are presented by other cells in combination with specific presentation molecules. Within T lymphocytes, several subsets – some of which (with later relevance to the digital immune systems) described below – carry distinct functional roles.

- *T helper cells* coordinate immune responses. T helper cells receive information from dendritic cells and other cells that present antigens and influence B lymphocytes, cytotoxic T cells and regulatory T cells through contact dependent signals and cytokines. Activation of T helper cells is a key step for the full development of adaptive immune responses.
- *Cytotoxic T cells* recognize antigens that indicate infection or transformation of body cells and induce death of such cells. Cytotoxic T cells therefore execute targeted elimination of cells that present specific antigens.
- *Regulatory T cells* limit the intensity and duration of immune responses. Regulatory T cells reduce activation of other T lymphocytes and B lymphocytes and help prevent excessive reactions and autoimmunity.
- *Memory T cells* arise from activated T lymphocytes after an immune response. Memory T cells persist for extended periods and can react rapidly to renewed exposure to the same antigen. Memory T cells therefore constitute the cellular part of immunological memory within adaptive immunity.

Adaptive immunity responds more slowly than innate immunity during the first encounter with an antigen but achieves high specificity and long-lasting memory through coordinated actions of B lymphocytes, T lymphocytes, antibodies, memory B cells, and memory T cells. In other words, the adaptive immunity achieves a form of biological sustainability and efficiency. By investing substantial resources during the initial encounter with an antigen and subsequently relying on highly specific memory cells, the adaptive immunity minimizes recurrent energetic and cellular costs while maintaining robust long-term protection.

C. Coordination and communication

The biological immune system continuously and bidirectionally interacts with the central nervous system. The central nervous system integrates information about the state of the organism and can influence immune responses, e.g. through hormones and autonomic regulation. Nerves provide communication links between peripheral tissues and the central nervous system, thereby affecting local immune processes. Antigen recognition as well as innate and adaptive responses therefore operate within a coordinated system that connects local immune mechanisms to central regulation at the level of the whole organism. Furthermore, a properly functioning immune system not only relies on accurate pathogen recognition, effective cell signaling, controlled inflammatory responses, and immune memory, but also on the finely tuned coordination of its components, as dysregulation may cause allergies or autoimmune diseases.

III. A REFERENCE ARCHITECTURE FOR DIGITAL IMMUNE SYSTEMS FOR CIVIL INFRASTRUCTURE

The reference architecture for digital immune systems aims to describe the building blocks and interactions that extend cyber-physical SHM systems for civil infrastructure, in terms of innate and adaptive defense capabilities. Civil infrastructure, including bridges, tunnels, and high-rise buildings, increasingly relies on cyber-physical SHM systems with dense sensor networks, embedded processing, communication links, and back-end services that operate as coupled systems comprising the physical structures and the SHM systems. In a coupled cyber-physical SHM system, threats no longer arise only from structural anomalies, but also from internal SHM system faults and from external cyberattacks. The proposed digital immune system operates on the digital side of this coupled system, while interacting with the physical structure through sensing and other SHM interfaces, without directly modifying the physical structure itself. The following subsections define threat categories and analyze the requirements that form the basis for the reference architecture (Fig. 2) and then describe the functional entities of the digital immune system, as summarized in Fig. 3. Finally, a formal specification of the reference architecture is provided that allows future developers to formally verify and rigorously assess whether a specific digital immune system implementation is structurally consistent with the threat categories addressed and with the architectural entities. The development of the reference architecture follows a design science research methodology (Peffers et al., 2007). Starting from the problem context and threat landscape, biological immune principles have been abstracted into functional roles and iteratively mapped to architectural entities of cyber-physical SHM systems, whereupon the resulting “artifact” (i.e., the reference architecture) has been evaluated against the requirements derived to ensure internal consistency and structural completeness.

A. THREAT CATEGORIZATION AND REQUIREMENTS ANALYSIS

The reference architecture satisfies several requirements that, so far, have received little attention in the design of SHM systems, such as (i) protection and monitoring of the coupled system of the physical structure and the SHM system, (ii) coverage of threat categories that include structural anomalies, SHM system faults, and cyberattacks, (iii) integration of rapid innate reactions and slower (but specific) adaptive reactions, (iv) clear distinction between global elements that act at system level and local modules that act at node level, (v) explicit support of learning and memory, enabling future reactions that build on past experience, and (vi) mechanisms that support regulation and safe return to a normal operating state after a threat has been handled. Potential threats can be categorized from two complementary viewpoints, a system-oriented view that distinguishes the physical structure and the SHM system, and a threat-oriented view that distinguishes internal and external threats, as illustrated in Fig. 2. In this case, "internal" and "external" are understood relative to the coupled system and the operational scope of the digital immune system rather than as statements about the ultimate physical cause of a threat. The reference architecture adopts the system-oriented view and therefore considers three threat categories:

- Threat category I (internal): Structural anomalies (including damage) within the physical structure
- Threat category II (internal): SHM system faults (including sensor faults) within the SHM system
- Threat category III (external): Cyberattacks (including sabotage) targeting the SHM system

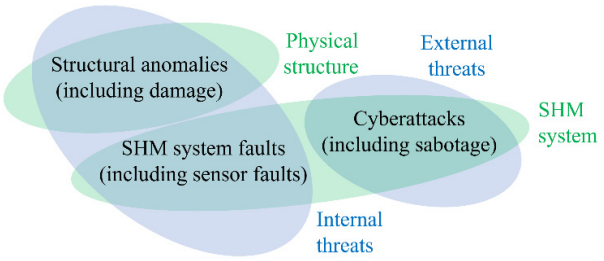

FIGURE 2. Threat categories relevant to digital immune systems for civil infrastructure.

Threat category III encompasses both generic intrusions into communication networks and stealthy integrity attacks targeting sensing, estimation, and decision-support functions. Evidence from adjacent critical-infrastructure domains, particularly IoT-based smart grids, shows that (i) state estimation can be compromised even when attacker knowledge is limited [26], (ii) machine-learning-based operational analytics are susceptible to evasion and poisoning attacks [27], (iii) hybrid encryption can enhance secure state estimation [28], and (iv) coordinated cyber-physical attacks can obscure manipulated control actions [29]. Collectively, this evidence illustrates the range of attack vectors and defensive strategies associated with threat category III. The insights are highly relevant to cyber-physical SHM, since transferable attack and defense patterns may contribute to the protection of tightly coupled cyber-physical SHM systems, used in critical infrastructure.

The conceptual distinction between *internal* and *external* threats, combined with the system-oriented separation between *physical structure* and *SHM system*, yields a two-dimensional categorization that logically comprises four possible threat categories. The fourth threat category refers to external threats acting directly on the physical structure, namely exogenous physical hazards in the broadest sense, including extreme environmental events and physical attacks, for example explosions. Hazards of this kind are not addressed directly by digital immune functions embedded in cyber-physical SHM systems, since these functions operate at the level of sensing, data processing, and system-level response. Instead, exogenous physical hazards become relevant to the digital immune system only through observable effects within the coupled system, namely as structural anomalies captured by category I or as SHM-system faults captured by category II, as discussed in [30]. In other words, normal environmental and operational conditions are not classified as threats by themselves, but form part of the operating context of the coupled system. Relevance to the digital immune system arises only when such conditions produce observable structural effects associated with category I or impair sensing, communication, or computation within the SHM system, thereby manifesting operationally as category II. Accordingly, category IV is retained for the sake of logical completeness, whereas the operative threat space of the proposed architecture is limited to categories I–III.

The threat categories are formally specified in a descriptive, architecture-oriented manner as follows. Let the sorts Threat, Subsystem, Origin, and Category be given. Let PS be a physical structure and SHM be a structural health monitoring system, where PS, SHM : Subsystem. Let int, ext : Origin. Let I, II, III, IV : Category.

$$\mathrm{PS} \neq \mathrm{SHM} \wedge \mathrm{int} \neq \mathrm{ext}.$$
$$\mathrm{I} \neq \mathrm{II} \wedge \mathrm{I} \neq \mathrm{III} \wedge \mathrm{I} \neq \mathrm{IV} \wedge \mathrm{II} \neq \mathrm{III} \wedge \mathrm{II} \neq \mathrm{IV} \wedge \mathrm{III} \neq \mathrm{IV}.$$

The following axioms restrict the sorts Subsystem, Origin, and Category to the explicitly introduced constants and thus ensure that no additional elements of these sorts are admitted within the scope of the present architectural specification.

$$\forall s : \mathrm{Subsystem}\ (s = \mathrm{PS} \vee s = \mathrm{SHM}).$$
$$\forall o : \mathrm{Origin}\ (o = \mathrm{int} \vee o = \mathrm{ext}).$$
$$\forall c : \mathrm{Category}\ (c = \mathrm{I} \vee c = \mathrm{II} \vee c = \mathrm{III} \vee c = \mathrm{IV}).$$

Let target : Threat → Subsystem and origin : Threat → Origin. The following exhaustiveness axiom ensures exhaustiveness of the case distinction:

$$\forall t : \text{Threat}\ ((\text{target}(t) = \text{PS} \vee \text{target}(t) = \text{SHM}) \wedge (\text{origin}(t) = \text{int} \vee \text{origin}(t) = \text{ext})).$$

The categorization function

$$\kappa : \text{Threat} \rightarrow \text{Category}$$

is defined by

$$\kappa(\text{t}) = \begin{cases} \text{I} & \text{if target}(t) = \text{PS} \wedge \text{origin}(t) = \text{int}, \\ \text{II} & \text{if target}(t) = \text{SHM} \wedge \text{origin}(t) = \text{int}, \\ \text{III} & \text{if target}(t) = \text{SHM} \wedge \text{origin}(t) = \text{ext}, \\ \text{IV} & \text{if target}(t) = \text{PS} \wedge \text{origin}(t) = \text{ext}. \end{cases} \quad (1)$$

As mentioned earlier, category IV is retained only for logical completeness but is not operationally relevant, as it is formally linked to category I by

$$\forall t : \text{Threat}\ (\kappa(t) = \text{IV} \rightarrow \exists t' : \text{Threat}\ ((\kappa(t') = \text{I}) \wedge (\text{target}(t') = \text{PS} \wedge \text{origin}(t') = \text{int}))).$$

The link is intentionally defined at the level of abstraction of the categorization and does not imply a direct event-level mapping between specific threats. Rather, the link reflects that exogenous physical hazards (category IV) lie outside the operational scope of the digital immune system and become relevant within the architecture only through internal structural effects represented by category I. The requirements and threat categorization defined above serve as structural design constraints for the development of the reference architecture. The abstraction and mapping of biological immune principles to digital entities in the following subsection are guided by the constraints formally specified above.

B. FUNCTIONAL ENTITIES OF THE DIGITAL IMMUNE SYSTEM

The functional design of the digital immune system follows the layered structure and component mapping introduced in the previous section. Following the traditional SHM setup, civil infrastructure being monitored is equipped with a cyber-physical SHM system that consists of distributed sensor nodes and a central server interconnected by communication links. Sensor nodes record data and provide embedded processing capacity close to the structure, whereas the central server aggregates information and executes high-level coordination and analysis. Digital immune systems, shown in Fig. 3 in the form of a reference architecture, can be realized as an additional functional layer on top of existing cyber-physical SHM systems. To avoid confusion, the functional entities of the digital immune system shown in Fig. 3 are not identical to sensor nodes or computers; instead, the components operate as global elements or as local modules that can be embedded into sensor nodes or installed on computers (or on both), depending on computational and communication constraints. Code migration between sensor nodes and the central server, for example as proposed in [31], allows the functional components to move at runtime and adapt their location to current requirements. The following subsections describe the functional roles of the global elements and the local modules in the innate layer and in the adaptive layer, as depicted in Fig. 3, followed by a short description of the system level coordination and deployment.

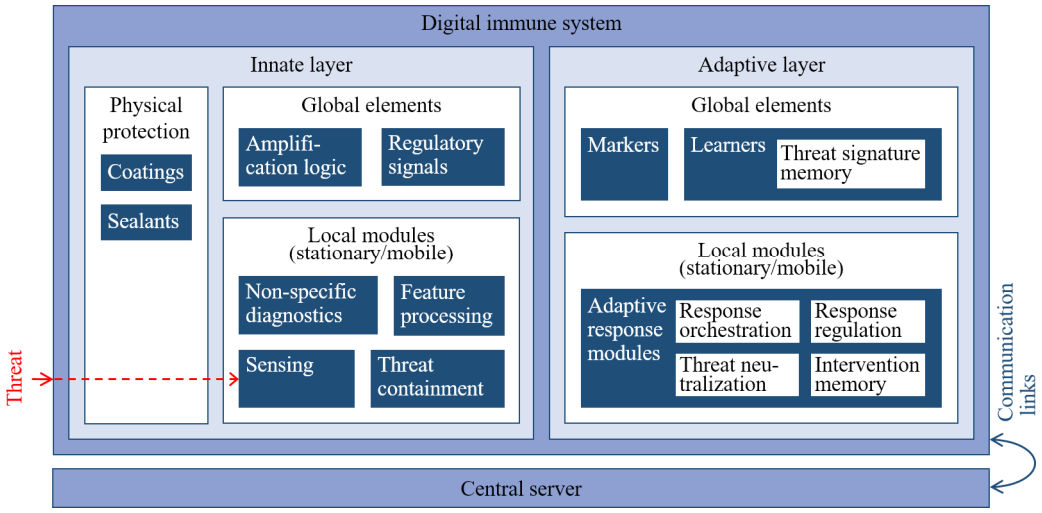

FIGURE 3. Reference architecture for digital immune systems for civil infrastructure.

1) INNATE LAYER

The innate layer forms the first internal defense level of the digital immune system. Components on this layer react quickly and in a largely non-specific way to all threat categories, i.e. structural anomalies, internal SHM system faults, and external cyberattacks. The innate layer comprises

physical protection as well as global elements and local modules, which together are intended to help prevent many threats from entering the coupled system of physical structure and SHM system and to enable fast system-wide reactions if threats are detected.

a) Physical protection

The physical protection is equivalent to the anatomical barriers of the biological immune system and constitutes the outer shell of the coupled system. In addition to protecting the physical structure, the shell reduces direct exposure of sensors and communication hardware, forming part of SHM systems, to environmental and mechanical impacts. Many potential threats may be prevented from entering the coupled system at this level and therefore may never appear as recognizable threats.

b) Global elements

Global elements of the innate layer operate in a system-wide manner and are functionally analogous to humoral components in the biological immune system. Together, amplification logic and regulatory signals transform local indications of anomalies into system wide responses and prepare the adaptive layer for focused analysis.

- *Amplification logic* (equivalent to the complement system) implements rule sets and algorithms that mark suspicious activity in data streams or communication channels and amplifies alerts originating from local modules. Amplification logic can, for example, increase the severity level of alerts that arise simultaneously from several sensors, aggregate evidence across different locations and escalate reactions when indications for structural anomalies, SHM system faults or cyberattacks accumulate.
- *Regulatory signals* (equivalent to cytokines) act as digital messages that coordinate reactions of local modules and adaptive components. Regulatory signals propagate information about ongoing events, such as "potential damage detected" or "sensor node offline", and adjust sampling rates, diagnostic depth or isolation levels across the SHM system.

c) Local modules

The local innate modules are typically deployed on stationary sensor nodes or on mobile sensor nodes (i.e., on mobile robots) and react close to the source of the data. The combination of sensing, non-specific diagnostics, feature processing, and threat containment, listed below, is intended to create a fast, robust defense level that operates continuously and can address structural anomalies, SHM system faults, and cyberattacks in a unified way, supplying the adaptive layer with preprocessed indicators for further analysis.

- *Sensing modules* (equivalent to pattern recognition receptors) capture raw measurements and status information from the physical structure and from the SHM system. The sensing module provides the data required for detecting structural anomalies, SHM system faults, and cyberattacks.
- *Non-specific diagnostic modules* (equivalent to macrophages) examine incoming data streams with simple – yet robust – checks that do not assume a detailed threat model. Non-specific diagnostic modules flag deviations that may originate from any of the three threat categories (structural anomalies, SHM system faults, cyberattacks) using, e.g., threshold tests, residual checks, or basic consistency checks between redundant sensors [32].
- *Feature processing modules* (equivalent to dendritic cells) derive compact descriptors from raw data indicating, e.g. mode shape changes, sensor drifts, or anomalous network traffic. Feature processing reduces data volume, filters noise, and prepares representations that can be evaluated efficiently by the adaptive layer.
- *Threat containment modules* (equivalent to natural killer cells) provide direct local reactions in the innate layer to limit potential damage. Typical actions include temporary isolation of faulty sensors or local blocking of suspicious network connections.

2) ADAPTIVE LAYER

The adaptive layer refines responses that originate from the innate layer, while both layers are mutually coupled through feedback loops. The adaptive layer learns from data, generates threat-specific markers and coordinates reactions that address structural anomalies, SHM system faults, and cyberattacks. Similar to the innate layer, the adaptive layer includes global elements and local modules.

a) Global elements

The global adaptive elements are analogous to the humoral adaptive components of the biological immune system and provide generation, storage, and management of markers at SHM system level. Markers, learners, and the threat signature memory form the global part of the adaptive layer and provide reference information for the adaptive response modules.

- *Markers* (equivalent to antibodies) are digital signatures and models that encode specific threat patterns in the data and support recognition of specific structural anomalies, SHM system faults, and cyberattacks.
- *Learners* (equivalent to B lymphocytes) are learning algorithms that derive new markers from labeled or otherwise confirmed events and adapt existing markers when threat characteristics change.
- *Threat signature memory* (equivalent to memory B cells) stores the set of markers together with previously learned knowledge and metadata, e.g. threat category, location, context, and performance indicators. The threat signature memory enables rapid reuse of markers and knowledge when similar situations occur and

maintains the marker repertoire over the lifetime of the monitored structure.

b) Local modules

Local adaptive response modules resemble the cellular part of the adaptive immunity of the biological immune system and translate system information into concrete system-level response actions within the coupled cyber-physical SHM system. More precisely, the local adaptive response modules are analogous to T lymphocyte subsets and execute decisions at node or node cluster level that combine information stemming from the global elements and from the local modules of the innate layer.

- *Response orchestration modules* (equivalent to T helper cells) receive information from the feature processing modules of the innate layer, select markers for the current context and decide which local and global reactions to activate. Typical tasks of response orchestration modules include choosing which learners to update, which containment actions to escalate and which diagnostics to schedule for suspected structural anomalies, SHM system faults, or cyberattacks.
- *Threat neutralization modules* (equivalent to cytotoxic T cells) carry out elimination actions at node or node cluster level, such as shutting down compromised nodes, permanently excluding corrupted data sources from structural assessments, or triggering secure reconfiguration of network routes and services.
- *Response regulation modules* (equivalent to regulatory T cells) limit the strength and duration of reactions by adjusting thresholds, de-escalating isolation levels, and suppressing unstable feedback loops between diagnostic modules. Response regulation therefore reduces false alarms and protects normal operation of the coupled system.
- *Intervention memory modules* (equivalent to memory T cells) store successful reaction sequences, including combinations of non-specific diagnostic checks, markers, containment actions and neutralization steps, linked to threat categories and contextual information. Future incidents may reuse the "intervention templates", which shortens reaction times and improves consistency of responses.

3) SYSTEM LEVEL COORDINATION AND DEPLOYMENT

The digital immune system operates within the coupled system of physical structure and SHM system. At system level, although individual sensor nodes act autonomously, the overall coordination relies on the central server and the communication links that interconnect sensors, local modules, global elements, and back-end services.

- The *central server* (equivalent to the central nervous system), besides performing the "traditional" SHM tasks, collects summaries from global elements and local modules, hosts coordination logic for the innate and adaptive layer, and can override or reinforce local decisions when the overall integrity of the coupled system is at risk.
- The *communication links* (equivalent to the nerves) connect local modules, global elements, and the central server across wired and wireless channels.

The specific deployment of global elements and local modules depends on the individual SHM system architecture. Some components may execute persistently on embedded sensor nodes, others on edge devices or in cloud back ends, and functions may migrate between the locations as computational or communication demands evolve. Independent from specific implementation choices, the reference architecture maintains a clear separation between innate and adaptive layers, between global elements and local modules, and between responses to structural anomalies, SHM system faults, and cyberattacks. The following section discusses integration options and research directions for embedding digital immune systems into existing and future SHM systems.

C. FORMAL SPECIFICATION OF THE REFERENCE ARCHITECTURE

This subsection presents the formal specification of the reference architecture using many-sorted first-order logic (Manzano, 1996). Rather than presenting a specific implementation, the formalization establishes an implementation-independent architectural definition that precisely characterizes the architecture and its entities, enabling a rigorous assessment of whether a concrete DIS implementation is consistent with the architectural principles and structurally consistent with the threat categories addressed. In addition, the formalization provides a solid basis for conducting formal checks and verification procedures in future DIS implementations, thus strengthening transparency, comparability, and reproducibility across different realizations of digital immune systems for civil infrastructure. The formal specification is organized into five parts, (i) the signature and layers, (ii) BIS-to-DIS mapping, (iii) roles and placement, (iv) architectural non-triviality, and (v) threat interaction. In the following formalization, the sort DISEntity refers exclusively to the digital functional entities of the reference architecture, i.e. local modules and global elements. Physical protection, communication links, and the central server are treated as environmental or infrastructural elements and are not included in DISEntity.

Signature and layers. Let the sorts BISComponent, DISEntity, and Layer be given. Let the sorts Threat and Category be given and let the function κ : Threat → Category (with the constants I, II, III, IV : Category) be as defined in Eq. (1). Let Innate, Adaptive : Layer.

$$(\text{Innate} \neq \text{Adaptive}) \\ \wedge (\forall \ell : \text{Layer}\ (\ell = \text{Innate} \vee \ell = \text{Adaptive})).$$

Let layer : DISEntity → Layer be a function. The predicates Sensing(·), NonSpecificDiagnostics(·), FeatureProcessing(·), ThreatContainment(·), AmplificationLogic(·), RegulatorySignal(·), ResponseOrchestration(·), ResponseRegulation(·), ThreatNeutralization(·), InterventionMemory(·), Marker(·), Learner(·), and ThreatSignatureMemory(·) are unary predicates over DISEntity. The predicates LocalModule(·) and GlobalElement(·) form a partition of DISEntity:

$$\forall e : \text{DISEntity}\ ((\text{LocalModule}(e) \lor \text{GlobalElement}(e)) \land \neg(\text{LocalModule(e)} \land \text{GlobalElement}(e))).$$

BIS-to-DIS mapping. Let map(·,·) be a binary predicate on BISComponent × DISEntity and assume

$$\forall e : \text{DISEntity}\ \exists c : \text{BISComponent}\ \text{map}(c,e).$$

The predicate map(*c*,*e*) captures structural-functional analogy rather than implementation equivalence. To make the BIS-to-DIS mapping formally meaningful, the mapping is constrained by role-consistency axioms that restrict the introduced BIS components to DIS entities with compatible architectural roles. Let PRR, Macrophage, DendriticCell, NaturalKillerCell, ComplementSystem, Cytokine, Antibody, BLymphocyte, MemoryBCell, THelperCell, CytotoxicTCell, RegulatoryTCell, and MemoryTCell be constants of sort BISComponent. The BIS-to-DIS mapping is constrained by the following role-consistency axioms:

$\forall e$: DISEntity (map(PRR, e) → Sensing(e))
$\forall e$: DISEntity (map(Macrophage, e) → NonSpecificDiagnostics(e))
$\forall e$: DISEntity (map(DendriticCell, e) → FeatureProcessing(e))
$\forall e$: DISEntity (map(NaturalKillerCell, e) → ThreatContainment(e))
$\forall e$: DISEntity (map(ComplementSystem, e) → AmplificationLogic(e))
$\forall e$: DISEntity (map(Cytokine, e) → RegulatorySignal(e))
$\forall e$: DISEntity (map(Antibody, e) → Marker(e))
$\forall e$: DISEntity (map(BLymphocyte, e) → Learner(e))
$\forall e$: DISEntity (map(MemoryBCell, e) → ThreatSignatureMemory(e))
$\forall e$: DISEntity (map(THelperCell, e) → ResponseOrchestration(e))
$\forall e$: DISEntity (map(CytotoxicTCell, e) → ThreatNeutralization(e))
$\forall e$: DISEntity (map(RegulatoryTCell, e) → ResponseRegulation(e))
$\forall e$: DISEntity (map(MemoryTCell, e) → InterventionMemory(e)).

The axioms constrain the mapping at the level of admissible functional roles and, in combination with the role and placement axioms introduced below, ensure the corresponding layer and placement consistency conditions without implying implementation equivalence.

Roles and placement. The role groups are defined as follows.

$$\begin{aligned}
R_{\text{IL}}(e) &:= (\text{Sensing}(e) \lor \text{NonSpecificDiagnostics}(e) \\
&\quad \lor \text{FeatureProcessing}(e) \lor \text{ThreatContainment}(e)), \\
R_{\text{IG}}(e) &:= (\text{AmplificationLogic}(e) \lor \text{RegulatorySignal}(e)), \\
R_{\text{AL}}(e) &:= (\text{ResponseOrchestration}(e) \\
&\quad \lor \text{ResponseRegulation}(e) \\
&\quad \lor \text{ThreatNeutralization}(e) \\
&\quad \lor \text{InterventionMemory}(e)), \\
R_{\text{AG}}(e) &:= (\text{Marker}(e) \lor \text{Learner}(e) \\
&\quad \lor \text{ThreatSignatureMemory}(e)).
\end{aligned}$$

Here, := denotes definitional abbreviation. The following axioms characterize admissible placement:

$$\forall e : \text{DISEntity} \left(\begin{array}{l} (R_{\text{IL}}(e) \rightarrow \\ (\text{LocalModule}(e) \land \text{layer}(e) = \text{Innate})) \\ \land\ (R_{\text{IG}}(e) \rightarrow \\ (\text{GlobalElement}(e) \land \text{layer}(e) = \text{Innate})) \\ \land\ (R_{\text{AL}}(e) \rightarrow \\ (\text{LocalModule}(e) \land \text{layer}(e) = \text{Adaptive})) \\ \land\ (R_{\text{AG}}(e) \rightarrow \\ (\text{GlobalElement}(e) \land \text{layer}(e) = \text{Adaptive})) \end{array} \right)$$

$$\forall e : \text{DISEntity} \left(\begin{array}{l} (\text{LocalModule}(e) \\ \quad \land\ \text{layer}(e) = \text{Innate}) \rightarrow R_{\text{IL}}(e) \\ \land\ (\text{GlobalElement}(e) \\ \quad \land\ \text{layer}(e) = \text{Innate}) \rightarrow R_{\text{IG}}(e) \\ \land\ (\text{LocalModule}(e) \\ \quad \land\ \text{layer}(e) = \text{Adaptive}) \rightarrow R_{\text{AL}}(e) \\ \land\ (\text{GlobalElement}(e) \\ \quad \land\ \text{layer}(e) = \text{Adaptive}) \rightarrow R_{\text{AG}}(e) \end{array} \right)$$

Architectural non-triviality. An architecture is considered non-trivial if it comprises at least one local and one global entity in each of the two architectural layers.

$\exists e$: DISEntity (LocalModule(e) ∧ layer(e) = Innate)
∧ $\exists e$: DISEntity (LocalModule(e) ∧ layer(e) = Adaptive)
∧ $\exists e$: DISEntity (GlobalElement(e) ∧ layer(e) = Innate)
∧ $\exists e$: DISEntity (GlobalElement(e) ∧ layer(e) = Adaptive).

Threat interaction. Let perceived(·,·) be a binary predicate on Threat × DISEntity. Perception is restricted to innate local sensing entities:

$$\forall t : \text{Threat}\ \forall e : \text{DISEntity}\ (\text{perceived}(t,e) \rightarrow (\text{Sensing}(e) \land \text{LocalModule}(e) \land \text{layer}(e) = \text{Innate})).$$

In summary, the architecture structurally situates threat perception within the innate-local sensing part of the system,

while the many-sorted first-order logic formalization serves as an implementation-independent specification layer for concrete DIS realizations. In software terms, sorts may be refined into typed software entities, functions and predicates into machine-processable operations and relations, and axioms into formally checkable constraints, thereby supporting architectural conformance checking, consistency verification, and query answering in DIS-oriented software systems or middleware. Accordingly, the formal specification is intended to assess whether a concrete DIS implementation is architecturally consistent with the proposed principles and complete with respect to the threat categories addressed. In this sense, the formalization defines structural constraints on layers, entity partitioning, admissible role placement, non-triviality, and threat interaction. By contrast, behavioral properties and performance criteria, such as detection accuracy, latency, false-alarm rates, and resilience gains, are not prescribed by the present formalization and remain to be addressed in implementation-specific validation studies

IV. RECOMMENDATIONS FOR IMPLEMENTING DIGITAL IMMUNE SYSTEMS FOR CIVIL INFRASTRUCTURE

Translating the reference architecture into operational cyber-physical SHM deployments requires more than a direct "one-to-one" coding of the components introduced in Section 3. In practice, digital immune systems for civil infrastructure must be instantiated in a way that the biological analogy remains functionally meaningful while still fitting heterogeneous SHM systems and the unique nature of civil infrastructure, legacy systems, and project-specific constraints. Rather than prescribing a single concrete implementation, this section outlines how the architectural building blocks can be instantiated in a generally applicable way, allowing that different SHM systems can realize their "own" digital immune systems on top of the same conceptual foundation. To structure the transition from architecture to implementation, the following subsection provides a process view that follows the immune-process logic from the biological immune system to its digital counterpart, with emphasis put on how recognition, amplification, response, memory, and regulation steps in the biological immune system can be mirrored by sequences of DIS operations. Thereupon, a functional view is provided that illuminates how the functional roles identified in the architecture, including global elements and local modules, can be deployed across concrete SHM components, such as stationary and mobile sensor nodes or central servers. To complement the implementation-oriented views, an illustrative use-case scenario is introduced, showcasing how representative architectural roles may be instantiated on an existing SHM test setup. From a software-engineering perspective, such realizations may be supported by typed metamodels, ontology-based representations, and middleware services that expose validation, consistency-checking, and reasoning functions through well-defined interfaces. This section concludes with general recommendations, intended to provide research and development paths required to mature digital immune systems for civil infrastructure, and the limitations of this study are summarized.

A. PROCESS VIEW

This subsection provides the process-oriented view that complements the reference architecture introduced earlier. The process view operationalizes the architectural entities defined in Section 3 and does not introduce additional structural assumptions. Rather than specifying a concrete implementation, the aim is to make explicit, which immune-like processes a cyber-physical SHM system must realize once a digital immune system is placed on top of it. Therefore, the canonical immune-process logic of the biological immune system is mirrored in a stepwise mapping. In the following list, the steps most relevant to the digital immune system are presented in two parts, (a) the biological immune process and (b) its digital counterpart. The process view serves as a reusable template for deriving DIS operation sequences that can be instantiated on cyber-physical SHM systems. The processes described below are visualized in Fig. 4, in which the mapping of the biological immune system components to the digital immune system is shown.

- ***Step 1: Anatomical barriers – PAMP-bearing pathogens***
  - BIS: The anatomical barriers (e.g., skin and mucous membranes) form the first level of innate immunity, shielding the body as physical protection and preventing a large fraction of pathogens and potentially harmful agents from invading the body.
  - DIS: The physical protection in the form of coatings and sealants constitutes the first level of the innate layer in the digital immune system for the coupled system of physical structure and SHM system.
- ***Step 2: PAMPs – Pattern recognition receptors/Complement system***
  - BIS: Pathogens and their conserved molecular patterns (e.g., PAMPs) that overcome the anatomical barriers are recognized by pattern recognition receptors on cells of the innate immunity and can also trigger the complement system as a humoral component. The combined recognition activates and amplifies cellular and humoral defense mechanisms and intensifies the ongoing innate immune response.
  - DIS: Threats and associated threat patterns that overcome physical protection are recognized within the innate layer by the sensing modules (equivalent to the pattern recognition receptors) (2a) and can independently activate the amplification logic (2b). Together, the dual recognition activates and amplifies the innate local and global defense in the digital immune system. Recognition at that stage is based on threshold-aware and uncertainty-informed criteria rather than on arbitrary deviations. Only deviations that exceed predefined or adaptive

thresholds are classified as suspicious patterns and passed on for further processing.

- ***Step 3: Pattern recognition receptors – Macrophages/Dendritic cells***
  - BIS: Activation of pattern recognition receptors on macrophages and dendritic cells brings the cellular components of innate immunity into an activated state. Macrophages increase phagocytosis and degradation of invading pathogen-derived material (bearing PAMPs), whereas dendritic cells take up pathogens and their products, extract characteristic features, and prepare the features as antigens for adaptive immunity.
  - DIS: Activation of the sensing modules causes the non-specific diagnostic modules and the feature processing modules to become active (representing local modules of the innate layer). Non-specific diagnostics locally examine structural and system data for anomalies, while feature processing extracts features from suspicious data fragments and prepares identified threat patterns for transfer to the adaptive layer.
- ***Step 4: Dendritic cells – T lymphocytes***
  - BIS: The dendritic cells form a functional link between innate immunity and adaptive immunity. The dendritic cells present processed antigen fragments to T lymphocytes and, in particular, activate T helper cells, thereby initiating the adaptive immune response and activating the cellular components of adaptive immunity.
  - DIS: The feature processing modules assume the abovementioned mediating role between the innate layer and the adaptive layer. The processed feature representations of threat patterns are transmitted via communication links to the adaptive response modules and activate, in particular, the response orchestration modules, i.e. the adaptive layer of the digital immune system is initialized and the adaptive immune response is initiated.
- ***Step 5: T helper cells – B lymphocytes/Cytotoxic T cells/Regulatory T cells/Memory T cells***
  - BIS: The activated T helper cells support and coordinate multiple cellular components within adaptive immunity. B lymphocytes with antibodies matching the specific antigen are stimulated to produce the specific antibodies and to form memory B cells; the cytotoxic T cells are activated for specific elimination of infected or transformed cells; the regulatory T cells are supported in their differentiation and function by the T helper cells and limit the strength of the response to prevent overreactions; and the memory T cells arise from a subset of activated effector cells and store successful response patterns for future encounters with the same antigen.
  - DIS: The response orchestration modules assume the central coordination function within the adaptive response modules. Depending on the deployment context, suitable realizations of adaptive functions include novelty detection and one-class classification methods for previously unseen threat patterns, incremental or online classifiers for continuous marker refinement, clustering-based discrimination for emerging pattern groups, sequence-learning methods for temporally evolving anomalies, and probabilistic or confidence-aware decision models for response regulation. To reduce false alarms and "digital allergies", learner functions should be combined with threshold adaptation, confidence estimation, uncertainty quantification, and confirmation by innate-layer evidence before strong adaptive actions are triggered.
- ***Step 6: B lymphocytes – Antibodies***
  - BIS: B lymphocytes activated by signals from the T helper cells proliferate and differentiate into cells that produce large quantities of specific antibodies against the recognized antigen; a fraction of the activated B lymphocytes differentiates into long-lived memory B cells that anchor antigen information in immunological memory.
  - DIS: In the digital immune system, activated learners generate specific markers in the form of digital signatures and models for the threat pattern recognized, based on information provided by response orchestration. Suitable learner classes include supervised or semi-supervised update mechanisms, incremental classifiers, novelty-detection models, and sequence-learning methods, whereas markers may take the form of digital signatures, probabilistic pattern models, or rule-based templates. To avoid unstable marker growth and overreaction, marker creation should be conditioned on confidence bounds, repeated evidence, and context consistency, so that only sufficiently supported threat representations enter the threat signature memory. At the same time, new entries are created in the threat signature memory, enabling fast and reliable recognition of known threats in the future.
- **Step 7 – Antibodies – Humoral components/Cellular effector cells**
  - BIS: Antibodies bound to antigen can activate the complement system (7a). Complement activation releases inflammatory mediators that act as regulatory signals and can increase cytokine release (7b). Together, the signals (i.e., complement-derived inflammatory mediators and cytokines) recruit immune cells (in particular macrophages) and enhance additional cellular effector mechanisms (e.g., macrophages and natural killer cells,

particularly when antibodies are bound to the target) at the site of antigen encounter (7c). The interaction of humoral components (antibodies, complement system, cytokines) and cellular effector cells (macrophages, natural killer cells) amplifies and focuses the immune response.

- DIS: Markers bound to their specific threat pattern activate the amplification logic as part of the global elements (7a). Activated amplification logic generates additional regulatory signals (7b) and stimulates the non-specific diagnostic modules to intensify inspection and direct the local modules toward affected data and system areas (7c). In parallel, the threat containment modules are directed toward suspicious entities to restrict or isolate the entities. The coordinated interaction of markers, amplification logic, and regulatory signals at the global level with non-specific diagnostics and threat containment at the local level amplifies and focuses the digital immune system response.

The following subsections build on this generic process logic by clarifying a minimal deployment view and by instantiating the process in an illustrative sensor-fault use-case scenario

B. FUNCTIONAL VIEW

Complementary to the process view described in the previous subsection, the functional view clarifies where the DIS functions may reside in a cyber-physical SHM deployment, i.e. how the functional roles introduced in the reference architecture can be instantiated. Importantly, the elements shown in the DIS reference architecture are functional roles rather than physical devices, i.e. a "sensor node" is not a DIS element itself, but an entity that may embed one or more local modules and may additionally execute selected global elements, depending on available resources and communication constraints.

Fig. 4 illustrates a minimal yet representative deployment scenario. Essentially, the deployment distinguishes between fast, non-specific first-line reactions in the innate layer and in-depth, context-dependent analysis and response in the adaptive layer. One or more *innate sensor nodes* (stationary or mobile) are deployed. The term "innate sensor node" denotes a sensor node hosting local modules of the innate layer (e.g., sensing, non-specific diagnostics, feature processing, threat containment). The innate sensor node(s) interface directly with the civil infrastructure by recording sensor data and by providing fast, non-specific first-line reactions. In parallel, one or more *adaptive sensor nodes* (stationary or mobile) are deployed. The term "adaptive sensor node" denotes a sensor node hosting local modules of the adaptive layer (e.g., response orchestration, response regulation, threat neutralization, intervention memory). Both innate and adaptive sensor nodes communicate with a *central computing unit* (CCU), typically through on-site computers or cloud-based backends.

The global elements – amplification logic and regulatory signals in the innate layer, as well as markers, learners, and the threat signature memory in the adaptive layer – are best understood as *software components* (more precisely, software services and data structures) with system-wide scope. In common SHM deployments, the global elements naturally reside on a CCU, but may also be deployed on sensor nodes or migrate across the system when needed. Such mobility does not change the conceptual distinction between global and local roles, i.e. "global" refers to the scope and availability of a function across the coupled system of physical structure and SHM system, not to a fixed physical location. For example, amplification logic can act as an event correlation and escalation service that aggregates evidence across nodes, while regulatory signals can be disseminated as coordination messages that steer sampling rates, diagnostic depth, and isolation levels across the SHM system. Similarly, markers represent digital signatures/models that must be accessible system-wide, while learners operate as training and update processes that generate or refine markers and maintain the threat signature memory. The following subsection illustrates the deployment logic by means of a representative use-case scenario instantiated by means of an existing SHM test setup.

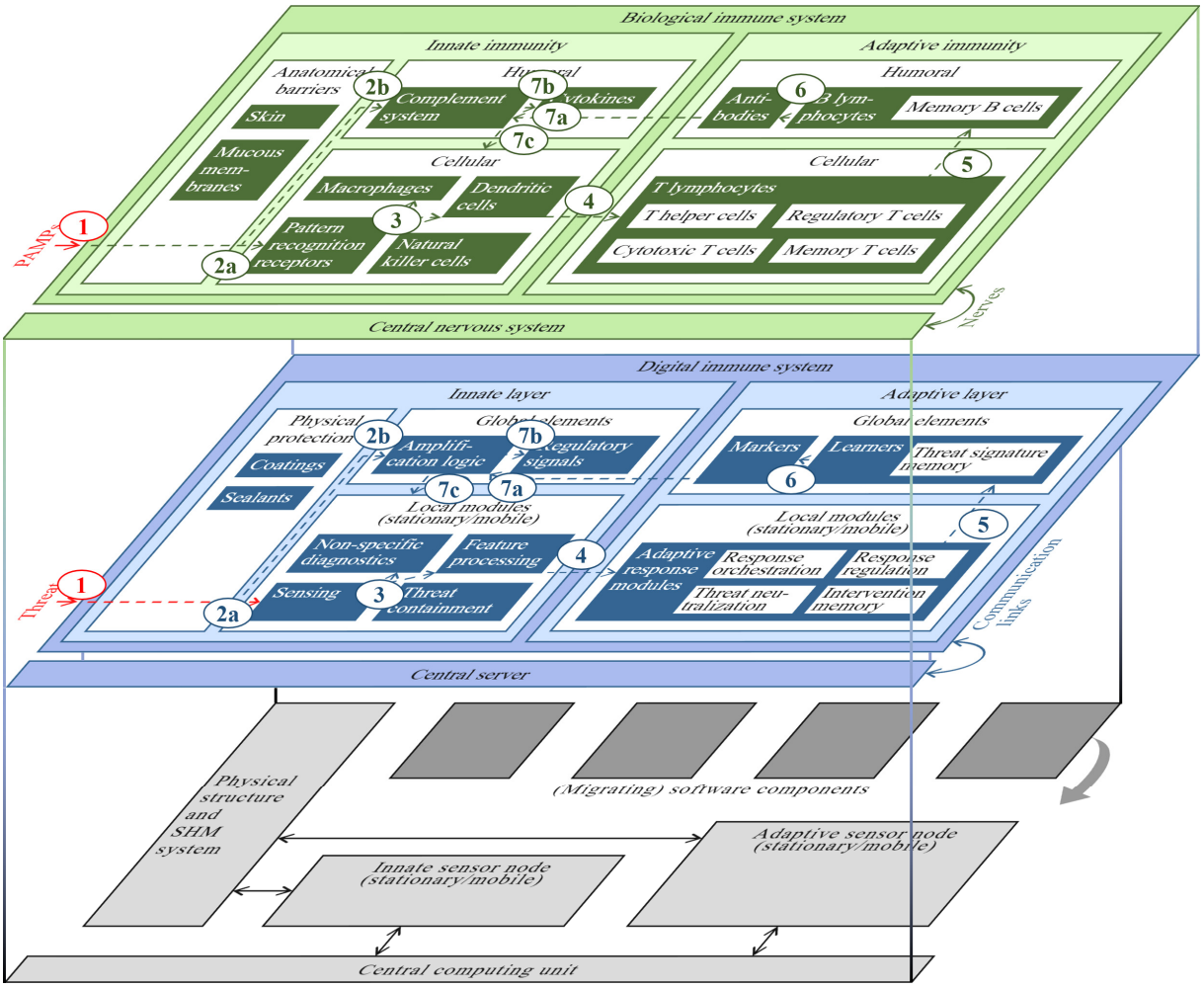

FIGURE 4. Process steps and mapping of the biological immune system (top) to the digital immune system (middle) and minimal deployment example (bottom).

C. ILLUSTRATIVE USE-CASE SCENARIO: SENSOR FAULT (THREAT CATEGORY II)

The following use-case scenario, aligned with the minimal deployment example illustrated in Fig. 4, describes how representative roles from the proposed reference architecture can be mapped onto a coupled system comprising the *physical structure and the SHM system*, *innate/adaptive sensor nodes*, *migrating software components*, and a *CCU*. The digital immune functions remain confined to the SHM system and its data-processing components, while the physical structure is addressed only through its observable responses. The use-case scenario serves an illustrative rather than prescriptive purpose, since the main contribution of the paper lies in a technology-agnostic reference architecture combined with an implementation-independent process logic. Algorithmic stacks, middleware choices, and deployment-specific integration strategies therefore remain intentionally open, given the substantial variation among SHM deployments for civil infrastructure with regard to sensing hardware, communication constraints, edge/cloud resources, and operational requirements. Nevertheless, a specific use-case scenario helps clarify how representative architectural roles may be realized, how the process logic unfolds, and how the architecture remains technically plausible without reducing the discussion to a system-specific implementation. The use-case scenario involves sensor-fault diagnosis, a common challenge in SHM, and is particularly relevant in light of the growing importance of reliable sensor-fault diagnosis in ageing and increasingly distributed SHM systems that increasingly rely on AI-supported structural assessment [33].

Fig. 5 and Fig. 6 show the laboratory SHM setup and the wireless sensor node, respectively, underlying the use-case scenario. The scenario is grounded in an existing laboratory SHM setup from prior sensor-fault-diagnosis research, and it is used here solely for illustrative architectural instantiation. Details on the hardware and software configuration are provided in [34] and [35]. The SHM system comprises four custom-made wireless sensor nodes equipped with microcontrollers and sensors, including accelerometers and environmental sensors, which communicate with the CCU responsible for bidirectional data transmission and data

management. The setup, originally developed and experimentally used in prior SHM research on decentralized sensor-fault diagnosis [35], i.e. not created specifically for the present use-case scenario, serves as an existing host platform onto which representative DIS roles can be mapped. Three nodes, indicated in Fig. 5, assume the role of innate sensor nodes, hosting sensing, non-specific diagnostics, feature processing, and threat containment, whereas a fourth node, functionally comparable to a cluster head, assumes the role of the adaptive sensor node. Role assignment is functional rather than hardware-inherent, i.e. role allocation is determined in software according to the architecture and to the use-case scenario rather than inferred from the node hardware. Furthermore, the architecture explicitly permits software functions to migrate between target nodes and resources of the CCU, allowing specialized analyses to be deployed only where needed rather than to reside permanently on every node. Accordingly, the architecture is intended to avoid a permanent additional resource burden across the SHM system and instead to realize selected DIS functions as resource-aware add-ons to existing modular SHM deployments After task completion, migrated software components remove themselves from the target nodes again, thereby releasing local resources, restoring the original software state, and reducing sustained communication and resource consumption, as demonstrated in earlier migration-based SHM research [31].

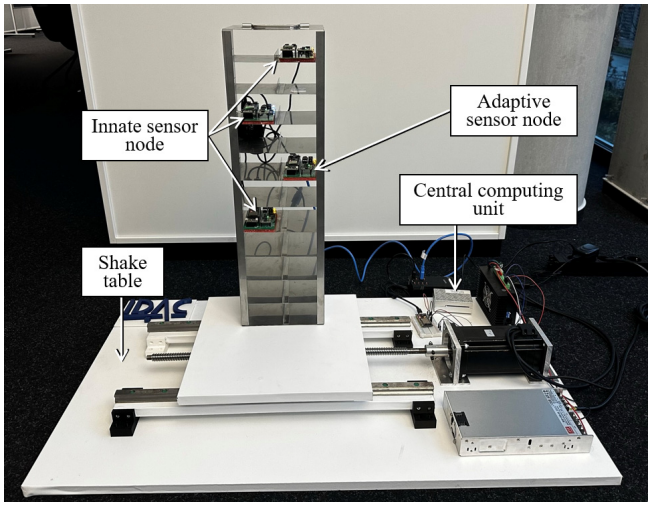

FIGURE 5. Laboratory SHM setup underlying the illustrative use-case scenario.

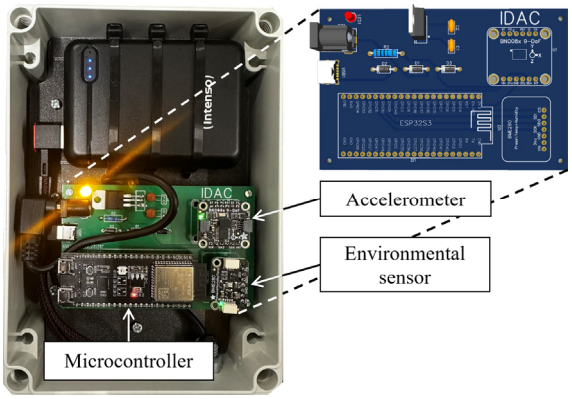

FIGURE 6. Custom-built wireless sensor node.

The sensor fault exemplarily considered in the use-case scenario is a bias fault affecting one accelerometer channel. Bias is selected because it represents a common and physically transparent type of sensor fault in SHM. In the use-case scenario considered herein, the bias corresponds to the laboratory setting reported in [34], in which a constant value is added to the acceleration response data collected by one wireless sensor node. The wireless sensor nodes and the CCU therefore provide the hardware and software environment in which representative roles of the DIS architecture are instantiated, including sensing, non-specific diagnostics, feature processing, threat containment, and adaptive response modules. Physical protection remains part of the innate layer as the outer protective boundary of the coupled system, while communication links connect the local modules of the innate layer and the adaptive layer on the sensor nodes with the global elements hosted on the CCU, functionally corresponding to the central server of the reference architecture. Fig. 7 illustrates an interaction sequence consistent with the BIS-to-DIS process steps introduced in Section IV.A. With physical protection remaining in place as the outer boundary of the coupled system (step 1, implicit in the present use case), the bias fault enters the operative threat space of threat category II. Once the sensor-fault footprint on the acceleration response data becomes observable, the respective abnormal sensor stream is recognized as a suspicious pattern (step 2). The classification is based on threshold-aware and uncertainty-informed criteria, ensuring that persistent low-level variations or noise do not trigger unnecessary responses. The abnormal sensor stream is perceived as suspicious by the sensing function of the innate sensor node, while an ensuing compact indication of anomaly may, in parallel, activate amplification logic on the CCU. The resulting local alert is thereby amplified without yet establishing a specific diagnosis. The suspicious pattern then activates modules for non-specific diagnostics and feature processing on the innate sensor node (step 3). Subsequently, the feature representation is transmitted to the adaptive sensor node (step 4), where response orchestration modules initiate the adaptive response and, if required, request additional

adaptive support for the context of the suspicious pattern. Adaptive support may be requested and deployed by means of migrating software components, for example following the migration-based SHM concept proposed in [31]. Depending on the adaptive support type, the adaptive stage may, for example, rely on neural-network-based analytical redundancy [33], classifiers using long short-term memory networks for complex sensor-fault patterns [37], or, where required, explainable AI (XAI) methods that support transparent interpretations of diagnostic decisions [38]. The bias-induced offset in the acceleration signal, constituting the suspicious pattern detected in the previous steps, is exemplified in Fig. 8, which shows the corresponding acceleration time series of the laboratory SHM system.

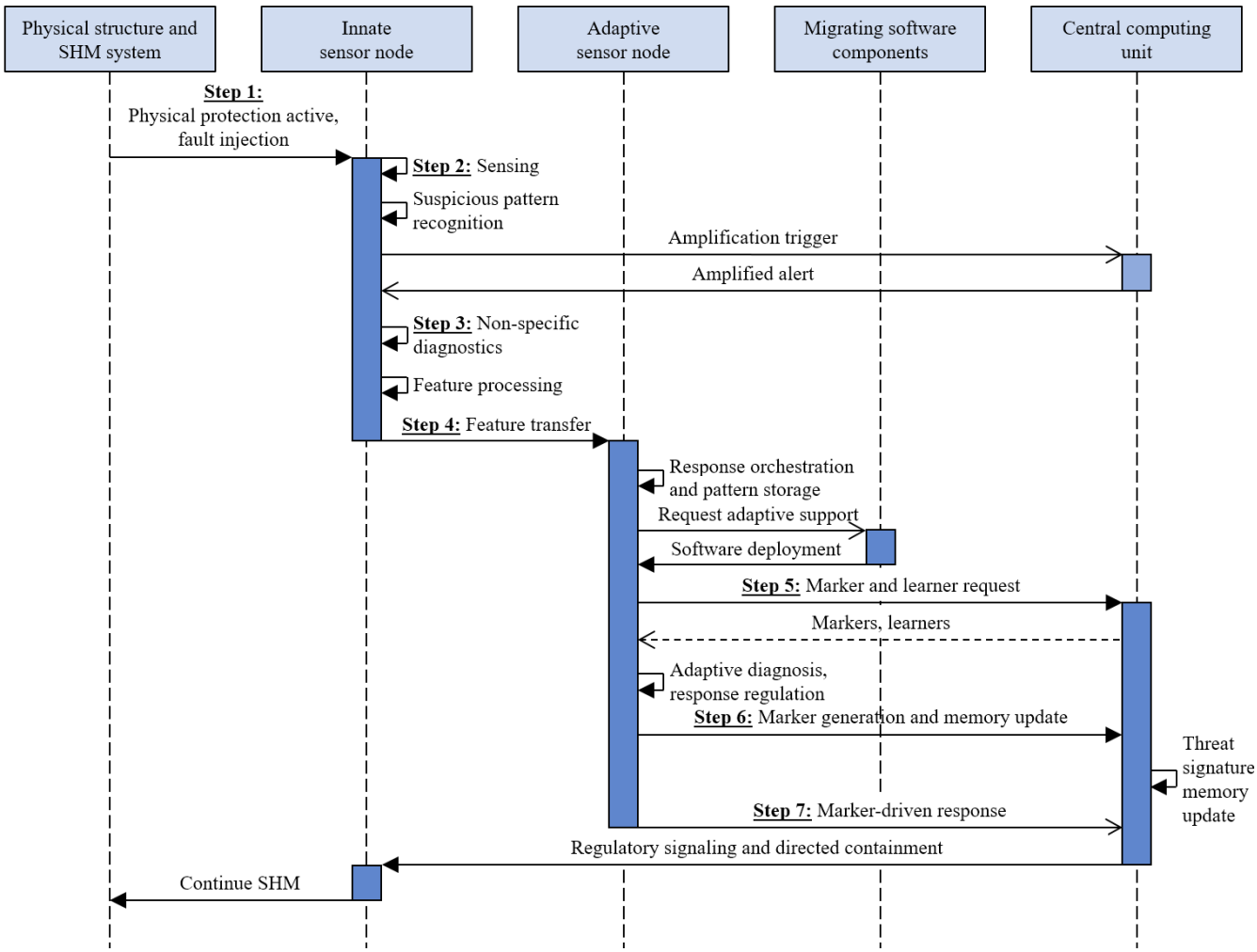

FIGURE 7. Interaction sequence of the illustrative use-case scenario.

Upon initiating the adaptive response, the adaptive sensor node requests access to markers and learners from the CCU to obtain historical context (step 5). The response may provide prior markers, access to learner functionality, and optional support for XAI. The adaptive diagnosis then evaluates whether the feature representation corresponds to a known bias-related sensor-fault signature or to a previously unseen case, and it regulates the response. Where necessary, Gradient × Input, Integrated Gradients, layer-wise relevance propagation, and Shapley additive explanations may be applied to analyze how sensor-level inputs contribute to the diagnosis, thus improving transparency, trust, and calibration of adaptive decisions [38]. To avoid unnecessary escalation, adaptive diagnosis should not rely on a single model output alone, but combine confidence estimates, repeated observations, and consistency with innate-layer evidence before generating new markers or triggering stronger neutralization actions. If a prior marker for the sensor fault is retrieved, it is reused; otherwise, the learner refines an existing marker or creates a new marker, while updating the threat signature memory on the CCU (step 6). Once the bias fault has been confirmed, the marker-driven response feeds back into the system-level logic (step 7). Marker-driven recognition activates amplification logic, which generates regulatory signals to guide the local response. On the innate side, local actions may include increased sampling, repeated checks, and

directed containment of the bias-affected channel. In parallel, response regulation limits overreaction, while threat neutralization remains available for additional, relatively strong interventions when necessary. A practically important outcome is fault accommodation, in which corrupted acceleration response data is isolated and replaced by virtual outputs derived from analytical redundancy. Fault accommodation essentially enables a self-healing-like continuation of SHM operation, since the monitoring function remains available despite the faulty sensor, following the concept proposed in [36]. Finally, the successful response pattern is consolidated as intervention memory on the adaptive sensor node, allowing similar sensor faults to be handled rapidly and consistently in future occurrences.

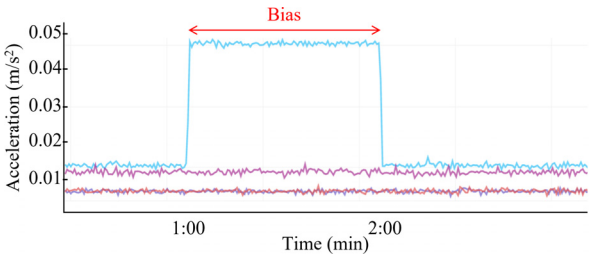

FIGURE 8. Bias fault in the illustrative use-case scenario.

Overall, the illustrative use-case scenario demonstrates in compact form how the proposed architectural elements and modules may interact in a bias-fault setting within a minimal SHM deployment. While the sequence remains implementation-open, the use-case scenario highlights how distinct algorithms, methods, and paradigms, including XAI, code migration, analytical redundancy, and self-healing, can be integrated into the proposed DIS process logic.

D. RECOMMENDATIONS AND LIMITATIONS

Building on the process and functional view presented in the previous subsections, this subsection summarizes general recommendations that delineate research and engineering steps required to mature digital immune systems from a reference architecture into robust, deployable technology for civil infrastructure. Furthermore, the limitations are summarized below, and it should be emphasized that the results of this study are conceptual in nature, i.e. no empirical or quantitative results are reported; the computational and communication overhead of DIS functions cannot be quantified within the scope of the present paper and is expected to depend strongly on implementation choices. A key lesson from biology is that immune protection is not only about strengthening responses, but also about regulating responses. Allergies and autoimmune diseases exemplify what happens when immune reactions are miscalibrated, i.e. the immune system overreacts to harmless stimuli or attacks host tissue. A direct analogue exists in cyber-physical SHM. If a digital immune system is insufficiently calibrated, it may overreact to benign operational variability, environmental changes, or gradual aging trends, causing persistent false alarms, unnecessary isolation of system entities, or disruptive interventions (“digital allergies”). In the extreme, poorly designed escalation logic or overly aggressive containment and neutralization actions may degrade monitoring performance or compromise the very system that the DIS is intended to protect. Accordingly, the following recommendations emphasize both effectiveness and safe regulation, outlining key research and engineering priorities for realizing digital immune systems in cyber-physical SHM deployments.

- *Immune-process modeling and analysis* should be further advanced, particularly with respect to the interaction between the innate and adaptive layers defined in Section 3, to improve the theoretical foundations of digital immune systems and to provide quantitatively validated design principles for future implementations.
- *Fault tolerance and cybersecurity* should be more closely tailored to the SHM domain, ensuring that fault and intrusion detection algorithms are more consistently aligned with the innate and adaptive defense layers of the digital immune system, which includes, e.g., protection against stealthy data-integrity attacks on estimation and learning pipelines, robustness of data analytics against adversarial perturbations, cross-layer consistency checks between commands and sensed effects, and secure estimation concepts based on cryptographic protection.
- *Simulation-based studies* of digital immune systems in realistic, large-scale cyber-physical SHM environments should be developed to validate architectural design choices, quantify resilience gains, and analyze emergent behavior under combined structural damage, fault, and cyberattack scenarios.
- *Data-centric workflows and artificial intelligence (AI) methods* should be integrated into the learners and markers of the digital immune systems, including uncertainty quantification and advanced numerical schemas, thus maintaining adaptive threat detection in a robust, explainable, and efficient form.
- *Memory-based intervention mechanisms* of the digital immune systems should be more rigorously formalized as reusable computational workflows, in which the intervention memory modules store parametrized response templates that can be automatically orchestrated across sensor nodes and back-end services when similar incidents reoccur.
- *Deployment strategies* for digital immune systems should be designed that embed the entities into existing cyber-physical SHM systems as modular, distributed services, explicitly leveraging parallel and distributed computing and edge-cloud integration, ensuring consistency with the layer separation and entity partition specified in Section III.C. At the same time,

deployment-specific trade-offs between resilience gains and additional computational, communication, and energy demand should be assessed explicitly. Related work in adjacent critical-infrastructure domains, such as smart grids, indicates that robustness-oriented cybersecurity functions may improve resilience but can also introduce non-negligible deployment-specific overhead.

- *Open, reproducible reference implementations* of key digital immune system entities, accompanied by benchmark datasets and standardized simulation scenarios, should be developed to enable practical validation as well as systematic comparison of alternative designs and to promote code and data sharing within the community.

Collectively, the aforementioned recommendations frame digital immune systems as a means that may help render modern cyber-physical SHM systems safer and to provide computational testbeds, in which high-performance and distributed computing, artificial intelligence, immunology, and SHM theory are interconnected. The recommendations are intended to help advance the maturity of digital immune systems for civil infrastructure and to foster interdisciplinary collaboration, thereby supporting the positioning of digital immune systems as a prominent application domain at the interface between civil engineering, computer science, and immunology.

Regarding the limitations of this work, it should be emphasized that the study is intentionally conceptual and is limited to providing a biologically inspired reference architecture and process logic that serves as a "blueprint" for practical implementations. Specific implementation choices (e.g., algorithms, training regimes, uncertainty handling, parameterization, and system integration) are intentionally left to the implementers and to the requirements of a given SHM deployment. Therefore, the manuscript does not report any implemented system, field deployment, or quantitative benchmark evaluation and, consequently, does not allow for quantitative effectiveness claims to be derived (e.g., false-alarm/false-negative rates, resilience gains, latency, or robustness under combined damage-fault-attack scenarios). The illustrative use-case scenario presented in Section IV is intended solely to clarify operational plausibility and architectural instantiation; it does not constitute a full DIS implementation or a separate validation study. Accordingly, the formal specification supports structural conformance checks for implementations with respect to the architectural principles and threat categories addressed, whereas behavioral and performance validation remain implementation-specific tasks. For example, while the architecture conceptually supports deployment across sensor nodes, edge devices, and central servers (including the potential migration of entities), system-specific constraints, such as bandwidth limitations, energy budgets, real-time deadlines, intermittent connectivity, and safety-certified operational requirements, are not explicitly addressed in detail. Further practical challenges include compatibility with existing middleware and back-end services, restricted access to installed hardware, and the need to avoid false alarms or disruptive overreaction when DIS functions are added to already operational SHM systems. Likewise, the importance of regulated interventions (to avoid "overreactions") is highlighted, while a formalized safety framework (e.g., verified escalation policies, fail-safe states, and bounded-response logic) is explicitly deferred to subsequent implementation and validation studies. Rather, simulation-based investigations, followed by prototype implementations and systematic validations, are positioned as the natural next step. Finally, given the complexity of immune processes, the biological-to-digital mapping is limited to those immune functions and architectural roles that admit direct digital counterparts in SHM of civil infrastructure, expressed in a form that is comprehensible to civil engineers.

V. SUMMARY AND CONCLUSIONS

Modern structural health monitoring systems form coupled cyber-physical systems, in which the physical structure and the SHM system must be protected as a whole. In the coupled system, threats include anomalies in the physical structure, internal SHM system faults, and external cyberattacks that may aim at sabotage or data theft, which motivates new SHM paradigms that have received limited attention so far.

This paper has proposed a biologically inspired reference architecture for digital immune systems for civil infrastructure, developed through a design science research process and formalized as a descriptive architectural specification. Rather than restricting the concept to a particular implementation, the architecture is formulated as a technology-agnostic and generally applicable framework that can be instantiated in diverse SHM contexts. The reference architecture, including the underlying threat categorization, has been formally specified in many-sorted first-order logic to ensure explicit separation of entity types, prevent category inconsistencies within the architectural specification, and enhance comprehensibility for civil engineers with limited background in formal logic. Following the biological immune system as conceptual model, the proposed architecture is organized into an innate layer and an adaptive layer. The innate layer comprises physical protection (e.g., coatings and sealants), local sensing, diagnostic, feature processing, and threat containment modules, together with global amplification logic and regulatory signals that are intended to distribute alerts throughout the system. The adaptive layer includes adaptive response modules (providing case-specific response orchestration, response regulation, threat neutralization and intervention memory) as well as global elements consisting of markers and learners that are devised to allow efficient adaptation to threats acting on the coupled system composed of physical structure and SHM system.

Based on the findings, recommendations for implementing digital immune systems into existing and future SHM systems

have been proposed to indicate potential future research directions. For example, future work may deepen immune-process modeling for digital immune systems and advance cybersecurity as well as fault diagnosis methods. In addition, reference implementations and benchmark scenarios may be devised to provide computational testbeds required to evaluate competing designs and to further position digital immune systems as an interdisciplinary research field integrating civil engineering, computer science, and immunology.

ACKNOWLEDGMENTS

The author would like to thank Professor Claudia Klümper (Hamm-Lippstadt University of Applied Sciences, Germany) as well as Kosmas Dragos, Thamer Al-Zuriqat, and Muhammad E. Ahmad (Hamburg University of Technology, Germany) for providing valuable technical expertise.

REFERENCES

[1] H. Sohn, C. R. Farrar, F. M. Hemez, D. D. Shunk, D. W. Stinemates, B. R. Nadler, and J. J. Czarnecki, A review of structural health monitoring literature: 1996–2001, Los Alamos Nat. Lab., Los Alamos, NM, USA, Rep. LA-13976-MS, 2002.

[2] G. Loubet, A. Takacs, and D. Dragomirescu, "Implementation of a battery-free wireless sensor for cyber-physical systems dedicated to structural health monitoring applications," *IEEE Access*, vol. 7, pp. 24679–24690, 2019.

[3] L. Cao, X. Jiang, Y. Zhao, S. Wang, D. You, and X. Xu, "A survey of network attacks on cyber-physical systems," *IEEE Access*, vol. 8, pp. 44219–44227, 2020.

[4] M. C. Montoya, C. E. Rubio-Medrano, and A. Kareem, "On the cybersecurity of smart structures under wind," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 251, Art. no. 105777, 2024.

[5] M. D. Champneys, A. Green, J. Morales, M. Silva, and D. Mascarenas, "On the vulnerability of data-driven structural health monitoring models to adversarial attack," *Structural Health Monitoring*, vol. 20, no. 4, pp. 1476–1493, 2021.

[6] Z. Zhang, R. Deng, D. K. Y. Yau, and P. Cheng, "Zero-Parameter-Information Data Integrity Attacks and Countermeasures in IoT-Based Smart Grid," *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6608–6623, 2021.

[7] Z. Zhang, M. Liu, M. Sun, R. Deng, P. Cheng, D. Niyato, M.-Y. Chow, and J. Chen, "Vulnerability of Machine Learning Approaches Applied in IoT-Based Smart Grid: A Review," *IEEE Internet of Things Journal*, vol. 11, no. 11, pp. 18951–18975, 2024.

[8] Z. Zhang, P. Cheng, J. Wu, and J. Chen, "Secure State Estimation Using Hybrid Homomorphic Encryption Scheme," *IEEE Transactions on Control Systems Technology*, vol. 29, no. 4, pp. 1704–1720, 2021.

[9] Z. Zhang, R. Deng, Y. Tian, P. Cheng, and J. Ma, "SPMA: Stealthy Physics-Manipulated Attack and Countermeasures in Cyber-Physical Smart Grid," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 581–596, 2023.

[10] C. R. Farrar and K. Worden, "An introduction to structural health monitoring," *Philosophical Transactions of the Royal Society A*, vol. 365, no. 1851, pp. 303–315, 2007.

[11] J. O. Kephart, G. B. Sorkin, and M. Swimmer, "An immune system for cyberspace," in *Proc. IEEE Int. Conf. Systems, Man, and Cybernetics: Computational Cybernetics and Simulation*, Orlando, FL, USA, Oct. 12, 1997.

[12] M. Swimmer, "Using the danger model of immune systems for distributed defense in modern data networks," *Computer Networks*, vol. 51, no. 5, pp. 1315–1333, 2007.

[13] R. Ghanea-Hercock, "Survival in cyberspace," Information Security Technical Report, vol. 12, no. 4, pp. 200–208, 2007.

[14] G. C. Silva, W. M. Caminhas, and R. M. Palhares, "Artificial immune systems applied to fault detection and isolation: A brief review of immune response-based approaches and a case study," *Applied Soft Computing*, vol. 57, pp. 118–131, 2017.

[15] S. Beck, "Immunogenomics: Towards a digital immune system," in Immunoinformatics: Bioinformatic Strategies for Better Understanding of Immune Function (Novartis Foundation Symposium 254), Chichester, United Kingdom: Wiley, 2003, pp. 223–230.

[16] J. Gómez-Gardeñes, P. Echenique, and Y. Moreno, "Immunization of real complex communication networks," *The European Physical Journal B*, vol. 49, no. 2, pp. 259–264, 2006.

[17] H. Yang, T. Li, X. Hu, F. Wang, and Y. Zou, "A survey of artificial immune system based intrusion detection," *The Scientific World Journal*, vol. 2014, Art. no. 156790, 2014.

[18] J. Timmis, T. Knight, L. N. de Castro, and E. Hart, "An overview of artificial immune systems," in *Computation in Cells and Tissues: Perspectives and Tools for Thought*, R. Paton, H. Bolouri, M. Holcombe, J. H. Parish, and R. Tateson, Eds., Berlin, Heidelberg, Germany: Springer, 2004, pp. 51–91.

[19] J. Gu, D. Lee, K. Sim, and S. Park, "An immunity-based security layer against Internet antigens," *IEICE Trans. Commun*., vol. E83-B, no. 11, pp. 2570–2575, 2000.

[20] U. Aickelin and S. Cayzer, "The danger theory and its application to artificial immune systems," in *Proc. 1st Int. Conf. Artificial Immune Systems*, Canterbury, United Kingdom, Sept. 11, 2002.

[21] J. Greensmith, "The dendritic cell algorithm," *Ph.D. dissertation*, School Comput. Sci., Univ. Nottingham, Nottingham, United Kingdom, 2007. [Online]. Available: http://ima.ac.uk/papers/greensmith_thesis.pdf. Accessed on: Dec. 10, 2025.

[22] Gartner, "What is a digital immune system and why does it matter?," Stamford, CT, USA: Gartner, Inc., 2022. [Online]. Available: https://www.gartner.com/en/articles/what-is-a-digital-immune-system-and-why-does-it-matter. Accessed on: Dec. 10, 2025.

[23] S. Kim, C. Hwang, and T. Lee, "Anomaly based unknown intrusion detection in endpoint environments," *Electronics*, vol. 9, no. 6, Art. no. 1022, 2020.

[24] 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.

[25] M. Manzano, *Extensions of First Order Logic*. Cambridge, United Kingdom: Cambridge Univ. Press, 1996.

[26] Z.-H. Yu and W.-L. Chin, "Blind False Data Injection Attack Using PCA Approximation Method in Smart Grid," *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1219–1226, 2015.

[27] J. Hao and Y. Tao, "Adversarial attacks on deep learning models in smart grids," *Energy Reports*, vol. 8, pp. 123–129, 2022.

[28] J. Wang, D. Shi, J. Chen, and C.-C. Liu, "Privacy-Preserving Hierarchical State Estimation in Untrustworthy Cloud Environments," *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1541–1551, 2021.

[29] J. Yang, G. Sun, and J. Yin, "Coordinated cyber-physical attack considering false overload of lines," *Protection and Control of Modern Power Systems*, vol. 7, Art. no. 44, 2022.

[30] K. Smarsly, "Digital immunization of civil infrastructure," in *Proc. 10th Civil Structural Health Monitoring Workshop*, Berlin, Germany, Sep. 6, 2026, submitted.

[31] K. Smarsly and K. H. Law, "A migration-based approach towards resource-efficient wireless structural health monitoring," *Advanced Engineering Informatics*, vol. 27, no. 4, pp. 625–635, 2013.

[32] K. Smarsly and K. H. Law, "Decentralized fault detection and isolation in wireless structural health monitoring systems using analytical redundancy," *Advances in Engineering Software*, vol. 73, pp. 1–10, 2014.

[33] H. Al-Nasser, T. Al-Zuriqat, K. Dragos, C. Chillón Geck, and K. Smarsly, "Identification of combined sensor faults in structural health monitoring systems," *Smart Materials and Structures*, vol. 33, no. 8, Art. no. 085026.

[34] C. Chillón Geck, T. Al-Zuriqat, M. Elmoursi, K. Dragos, and K. Smarsly, "AIoT-enabled decentralized sensor fault diagnosis for structural health monitoring," in *Proc. 11th European Workshop on Structural Health Monitoring (EWSHM 2024)*, Potsdam, Germany, Jun. 10, 2024.

[35] T. Al-Zuriqat, C. Chillón Geck, K. Dragos, G. D. Manolis, and K. Smarsly, "Decentralized sensor fault diagnosis for wireless structural health monitoring systems using Artificial Intelligence of Things," in *Proc. 15th International Workshop on Structural Health Monitoring (IWSHM 2025)*, Stanford, CA, USA, Sep. 9, 2025.

[36] K. Smarsly and D. Hartmann, "AMBOS – A self-managing system for monitoring civil engineering structures," in *Proc. XVI Workshop on Intelligent Computing in Engineering (EG-ICE 2009)*, Berlin, Germany, Jul. 15, 2009.

KAY SMARSLY is Professor of Civil and Environmental Engineering at Hamburg University of Technology and founding director of the Institute of Digital and Autonomous Construction. He earned his doctorate in civil engineering from Ruhr University Bochum in 2008 with summa cum laude. From 2010 to 2013, he was a DFG Research Fellow at Stanford University, and he has also conducted research at Carnegie Mellon University and held a fellowship at Berlin University of Technology. In 2013, he was appointed Professor and Chair of Computing in Civil Engineering at Bauhaus University Weimar and, in 2021, accepted the appointment at Hamburg University of Technology. He received the Hamburg Teaching Award in 2023, and he was elected to the North Rhine-Westphalia Academy for Sciences, Humanities and the Arts in 2024.

Professor Smarsly is Chair of the German Association of Computing in Civil Engineering (GACCE) and member of the Board of Directors of the International Society of Computing in Civil and Building Engineering (ISCCBE). He is also editorial board member and active reviewer for numerous international journals and research funding agencies and has supervised more than 250 doctoral, master's, and bachelor's theses as well as student projects. His research, focusing on digital twins, construction robotics, building information modeling, smart monitoring, and artificial intelligence in civil engineering, is documented in over 250 publications.

[37] H. Al-Nasser, K. Dragos, and K. Smarsly, "Explainable sensor fault diagnosis for structural health monitoring," in *Proc. 11th ECCOMAS Thematic Conference on Smart Structures and Materials (SMART 2025)*, Linz, Austria, Jul. 1, 2025.

[38] H. Al-Nasser, K. Dragos, and K. Smarsly, "Comparative analysis of explainable artificial intelligence methods in sensor fault diagnosis for vibration-based structural health monitoring," in *Proc. XIII International Conference on Structural Dynamics (EURODYN 2026)*, Hannover, Germany, Sep. 27, 2026 (submitted).