information

Article

Digital-Twin-Based Security Analytics for the Internet of Things

Philip Empl * and Günther Pernul

Chair of Information Systems, University of Regensburg, 93053 Regensburg, Germany
* Correspondence: philip.empl@ur.de

Abstract: Although there are numerous advantages of the IoT in industrial use, there are also some security problems, such as insecure supply chains or vulnerabilities. These lead to a threatening security posture in organizations. Security analytics is a collection of capabilities and technologies systematically processing and analyzing data to detect or predict threats and imminent incidents. As digital twins improve knowledge generation and sharing, they are an ideal foundation for security analytics in the IoT. Digital twins map physical assets to their respective virtual counterparts along the lifecycle. They leverage the connection between the physical and virtual environments and manage semantics, i.e., ontologies, functional relationships, and behavioral models. This paper presents the DT2SA model that aligns security analytics with digital twins to generate shareable cybersecurity knowledge. The model relies on a formal model resulting from previously defined requirements. We validated the DT2SA model with a microservice architecture called TWINSIGHT, which is publicly available, open-source, and based on a real industry project. The results highlight challenges and strategies for leveraging cybersecurity knowledge in IoT using digital twins.

Keywords: digital twin; security analytics; internet of things; cybersecurity

check for updates

Citation: Empl, P.; Pernul, G. Digital-Twin-Based Security Analytics for the Internet of Things. *Information* **2023**, *14*, 95. https://doi.org/10.3390/info14020095

Academic Editors: Libing Wu and Krzysztof Szczypiorski

Received: 30 November 2022
Revised: 24 January 2023
Accepted: 30 January 2023
Published: 4 February 2023

1. Introduction

Organizations utilize emerging technologies around the Internet of Things (IoT) to stay competitive and build knowledge. In the Industrial Internet of Things (Industrial IoT), sensors complement existing systems to enact more data leading to more accurate predictive maintenance. Thereby, intertwining two complementary worlds is crucial in enabling IoT use cases. Most commonly, information technology (IT) systems are connected to operational technology (OT), processing volumes of data to gain new knowledge about machines. This intertwining is a tremendous cybersecurity challenge for organizations, as the OT remains reachable via the internet. Additionally, the myriad of lifecycle participants, devices, and systems creates an opaque and complex scenario. Knowing the IoT systems well is necessary for efficient cybersecurity management.

Lifecycle-centric cybersecurity management is crucial. Most recently, supply-chain attacks have been emerging, compromising the chain's weakest link. The ENISA states that half of the attacks are attributed to advanced persistent threats (APT) groups, mainly accessing data [1]. Consider a manufacturer of machines, i.e., cyber-physical systems, vending IoT assets built out of physical components from suppliers to business customers. Business customers own IoT assets and operate, maintain, and recycle IoT assets through external service providers. From this example, different lifecycle participants interact with an IoT system, each opening up new attack vectors. Nevertheless, attacks are not limited to supply-chain attacks. Many well-known attacks exist, e.g., Sybil attacks, wormhole attacks, ransomware attacks, and distributed denial service attacks [2]. These attacks can take place throughout the lifecycle of an IoT system, and they are getting more advanced and harder to detect.

Security analytics is a paradigm to enhance cybersecurity in the Industrial IoT. It applies big data analytics techniques aggregating and internalizing external cybersecurity knowledge [3]. Security analytics collects data from various sources and allows the correlation of data, e.g., whether an incident happened or is even likely to happen. Security analytics is essential to organizations, but the less efficient the models, the less efficient the security analytics. Nevertheless, APTs urge organizations to rely on internal and external knowledge covering the unforeseeable as well as possible. Thus, generating knowledge needs to come along with sharing. Sharing knowledge between lifecycle participants about threats and incidents improves cybersecurity [4,5]. For instance, over 90% organizations state that they rely on actionable cybersecurity knowledge from third parties and partners [6]. As generating valuable knowledge to be shared is challenging, organizations deploying complex and lifecycle-centric IoT systems require different and novel approaches.

The Industrial IoT requires more cooperation and collaboration between lifecycle participants (see ISO 27036) and more advanced techniques to cope with sophisticated cybersecurity attacks. The industry has long used the digital twin paradigm to map a physical asset to its virtual representation through the lifecycle by replicating or simulating the IoT [7]. Its use is not limited to operational scenarios, e.g., digital twins assist security testing, enhance cyber situational awareness, and improve intrusion detection systems [8,9]. Estimates say that 80% of organizations instrument some forms of digital twins for cybersecurity scenarios, whereby 85% of security officers agree that digital twins unleash even more efficient detection and mitigation [10]. By abstracting physical assets, digital twins reduce the complexity of problem solving and provide a semantic layer to the virtual representations. Thereby, they ensure a more detailed analysis of the physical asset's state and historical data, leveraging security analytics on top of the data structure of digital twins.

Digital twins show potential in knowledge generation (coupled with security analytics) and knowledge sharing. By bringing security analytics and digital twins together, we aim to enhance cybersecurity in the Industrial IoT along the lifecycle and push cybersecurity knowledge-sharing research. We address the research gap that no unified framework for digital-twin-based security exists and add a replication-based intrusion detection approach to it. This will help future research in the development of digital-twin-based security analytics. The following research question guides this paper: *"How can one align security analytics and digital twins?"* Recent research has already examined digital twins and security analytics specifically but not as a whole. By aligning security analytics with digital twins, this paper contributes to a more secure Industrial IoT by generating cybersecurity knowledge and demonstrating how to share this knowledge throughout the lifecycle. Our contributions are summarized in the following:

1. We comprehensively align security analytics with digital twins and illustrate how to generate and share cybersecurity knowledge between lifecycle participants.
2. We provide a novel formal model for digital twins and security analytics. This formal model assists in implementing digital-twin-based security analytics use cases.
3. We envision the DT2SA model for digital twins and security analytics. This model integrates the Industrial IoT and mediates a global understanding for further research and practical adoption.
4. We instantiate the DT2SA model by implementing a microservice architecture leveraging digital twin-based security analytics based on a real-world research project. TWINSIGHT enables digital-twin-based threat and incident detection using open-source software.

This paper is structured as follows. Section 2 provides the fundamental background knowledge on digital twins and security analytics. We further discuss related research. In Section 3, we conceptually elaborate on knowledge generation and sharing in the cybersecurity domain resulting in requirements. Section 4 takes up these requirements and puts entities and their relationships shaping a formal model. While Section 4.3 outlines the DT2SA model based on the formal model, Section 6 validates the DT2SA model concerning the requirements and introduces TWINSIGHT, a digital-twin-based security analytics microservice architecture. Section 7 concludes the paper and highlights future research.

2. Background and Related Work

In the following sections, we present relevant background on digital twins for security operations in Section 2.1 and for security analytics in Section 2.2. Section 2.3 highlights the related work and existing research focusing on digital twins, security analytics, and knowledge generation and sharing.

2.1. Digital Twins for Security Operations

The digital twin paradigm is deeply grounded in the Industrial IoT, representing a physical asset (e.g., an entity, system, process, or person) that is mapped throughout its lifecycle to a virtual counterpart built on semantics [7]. The digital twin relies on descriptive and dynamic asset data, dynamic environment data, historical asset and dynamic data, and semantics [11]. The application scenarios are diverse and are not limited to operational use. The application of digital twins for cybersecurity scenarios is a trend in research, e.g., Lopez et al. are developing an authorization mechanism through digital twins in 5G environments. Technically, digital twins are executed via so-called operational modes, i.e., simulation, analysis, and replication, which differ as follows:

- *Analytics*—using state data with statistical analysis.
- *Simulation*—using specification data with emulation or simulation techniques.
- *Replication*—using specification and state data with emulation or stimuli techniques.

These operation modes serve different use cases and enable various application scenarios, e.g., intrusion detection, security testing, security training, or penetration testing [8]. Digital twins are also considered an additional layer of security for the IoT that manages incident response [12]. Of course, digital twins present new cybersecurity challenges, but this has already been addressed by research [13]. However, we treat digital twins as an additional layer for more efficient security analytics.

2.2. Security Analytics

In the era of big data, new technologies are emerging to support efficient real-time data processing and analysis, which has led to the term big data analytics [14]. Of course, the need for big data processing in IoT is obvious and key to dealing with the vast amount of heterogeneous IoT data. Big data processing technologies provide a widely used foundation for further analysis. Siow et al. [15] provided an excellent summary of big data analytics in IoT and defined the following five analytical operations:

- Descriptive analytics: *What has happened?*
- Diagnostic analytics: *Why did it happen?*
- Discovery analytics: *What is happening?*
- Predictive analytics: *What will happen?*
- Prescriptive analytics: *What should one do?*

Descriptive or diagnostic analytics provides hindsight, discovery analytics provides insight, and predictive or prescriptive analytics provides foresight. From a cybersecurity perspective, big data analytics is crucial.

Due to the large amount of (un)structured security-relevant data, traditional security information and event management (SIEM) systems are reaching their performance limits [16]. Security analytics is concerned with applying big data processing technologies to cybersecurity and describes the aggregation and analysis of security-relevant data [3]. However, there is no clear definition of security analytics. We define security analytics as follows:

> Security Analytics is a repertoire of capabilities and technologies for the systematic processing and analysis of data to identify threats and imminent incidents.

We see security analytics as an evolution of SIEM with additional operations, e.g., intrusion detection systems, behavioral or network analysis [17], and knowledge sharing.

2.3. Related Work

Big analytics and security analytics.is trending. Ackoff [18] defined the DIKW and presented how to generate wisdom from data. Siow et al. [15] summarized five analytic operations for big data analytics and their alignment with the DIKW. These operations could be anchored in the security analytics domain [3] and were already aligned with security analytics [19]. We go beyond these analytic operations to illustrate how cybersecurity knowledge can be generated by linking operations, expertise, and digital twins.

Generating and sharing cybersecurity knowledge has already been addressed. The incident response process shows data, observables, indicators of compromise, and incidents [20]. Böhm et al. [4] elaborated on knowledge transformation in security analytics and formalized different types of knowledge, i.e., explicit and implicit knowledge. We aim to improve knowledge generation and share research by linking security analytics with digital twins to promote knowledge generation in cybersecurity.

Eckhart et al. have developed CPS Twinning and CPS Replication, both frameworks for creating and deploying digital twins for cybersecurity scenarios [21]. Digital twins for security operations can be equipped with various capabilities, such as analytics, penetration testing, and intrusion detection. For example, Dietz et al. [22] integrated digital twin security simulations into a security operations center to assist analysts with security testing and monitoring IoT assets. Damjanovic-Behrendt [23] defined a microservice architecture of the digital twin that refers to security analytics as data analytics. Other service management tasks, such as incident detection and responding, complement security analytics. Since some research already presented digital twins for security analytics, we comprehensively go beyond this understanding and formalize security analytics and related knowledge sharing with digital twins. In doing so, we select an application scenario to demonstrate our overall model.

In summary, the existing literature does not comprehensively address security analytics in IoT. While there are approaches to align particular analytical operations with digital twins, a comprehensive view still needs to be provided. Our goal is to establish a comprehensive model and bring the previously practice-oriented security analytics into the realm of science. In this way, we will create a balance and improve the use of cybersecurity knowledge in IoT.

3. Managing Cybersecurity Knowledge

The following sections make steps towards the formal model. We first describe the generation of knowledge in Section 3.1. After, we illustrate knowledge sharing between lifecycle participants interacting with digital twins in Section 3.2. We then describe requirements for the formal model in Section 3.3, resulting in the intertwining of environments in Section 3.4.

3.1. Cybersecurity Knowledge Generation

Without knowledge, there is nothing to share. As already mentioned, knowledge is a product of information [18], whereby knowledge generation requires human interaction. In doing so, humans explore data, find insights, formulate hypotheses, and generate knowledge repetitively (sensemaking loop) [24]. However, generating knowledge is challenging and requires interaction with data and models. Research and industry cope with analytical operations by describing the appropriate mix of technologies, techniques, and cognitive abilities. The literature summarizes these analytical operations as descriptive, diagnostic, detective, predictive, and prescriptive ones [15]. We analyzed blog posts from several large organizations, i.e., IBM and Microsoft, confirming this notion but not all covering detective/discovery analytics. However, these five operations ensure knowledge generation all throughout the process.

Cybersecurity is also concerned with analytics. For example, there is an overlap between big data analytics and security analytics for deriving patterns for incident detection. There are similarities between the incident response process [20] and the DIKW [18]. Aside from the naming, the data, observables, indicators of compromise, and incidents share the same relationships as the DIKW. We summarize and transfer these concepts into the cybersecurity context. Figure 1 shows our approach to generating cybersecurity knowledge. Descriptive operations contextualize data into observables that describe specific events within an attack. Diagnostic operations involve analysis and correlation of historical observables, i.e., forensic investigation. Detective operations are suitable for detecting incidents based on indicators of compromise in real-time, e.g., intrusion detection systems. Predictive and prescriptive analytic operations are used to predict incidents and derive actions.

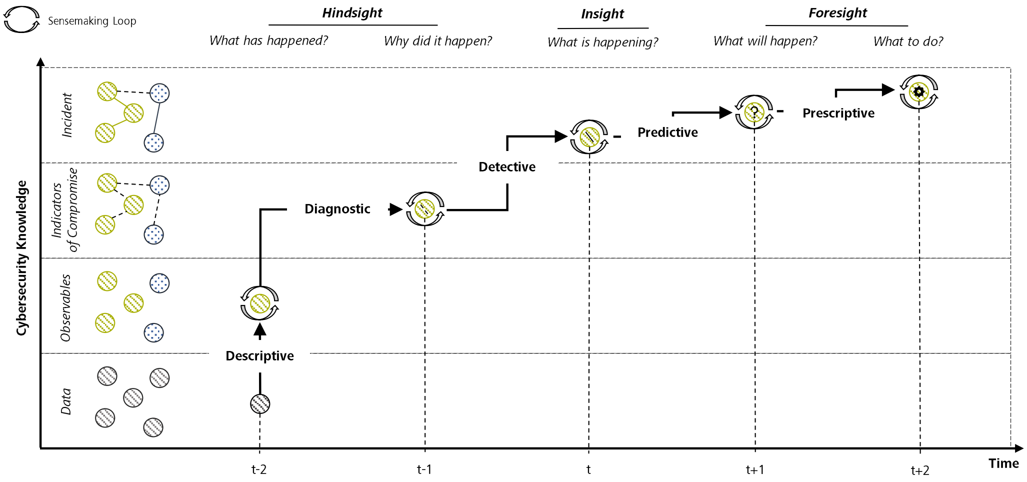


Figure 1. Cybersecurity knowledge generation.

In practice, analysts elaborate observables from many unstructured data sources (i.e., IP addresses). Malicious IP addresses represent so-called indicators of compromise. When they occur in a particular combination with other indicators of compromise, we refer to them as incidents. Nevertheless, these operations are part of the sensemaking loop, since knowledge is shaped by human interaction. Now that we know how knowledge is created in cybersecurity, we examine knowledge sharing.

3.2. Cybersecurity Knowledge Sharing

Knowledge sharing is critical in the Industrial IoT because it is a playground for different players and technologies. It includes many heterogeneous devices, a variety of communication protocols, and standards under development. From a manufacturing perspective, the collaboration between supply chain participants is broader than services and systems. Data must also be available across organizational boundaries. Digital twins have emerged as a paradigm to meet management within the circular economy [25]. They have proven their strength in sharing knowledge with lifecycle participants as long as communication channels are kept secure [26].

Collaboration between lifecycle participants must be strengthened, as this will improve observables and indicators of compromise through external knowledge [4]. For example, the greater the knowledge about indicators of compromise, the more efficient security analytics. To promote collaboration among lifecycle stakeholders, we must first understand their roles in the Industrial IoT. We summarize the leading roles in Industrial IoT: manufacturer, supplier, distributor, maintainer, and owner (see Figure 2). Since digital twins promote knowledge sharing in the Industrial IoT, we illustrate their bidirectional communication with their physical counterparts. A physical asset consists of several components from multiple suppliers assembled by a manufacturer. A distributor brokers the asset to an owner, who contracts a maintainer to provide services and maintenance. Digital twins communicate bidirectionally and serve as a single point of truth, eliminating information asymmetry. For example, they notify an owner if an incident occurs or report vulnerabilities to a supplier or manufacturer. Given this knowledge generation and sharing notion, we can define the key requirements.

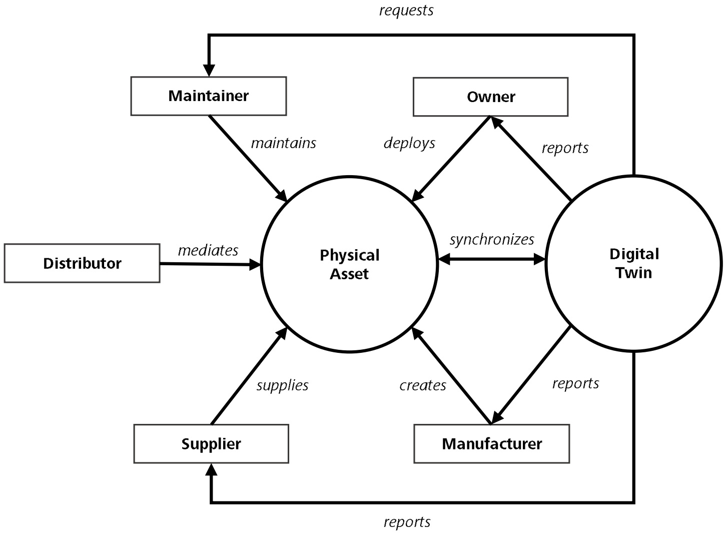


Figure 2. Cybersecurity knowledge sharing.

3.3. Requirements

Aligning security analytics and digital twins requires determining requirements concerning digital twins, security analytics, and knowledge sharing. We define the requirements as follows:

REQUIREMENT 1 (DIGITAL TWIN). The digital twin comprises a physical, a virtual, and a communication component [7,27]. It describes descriptive, dynamic, environmental, historical, and semantical data [11]. For cybersecurity operations, the digital twin operates in simulation, analytics, and replication modes [21,28].

REQUIREMENT 2 (SECURITY ANALYTICS). Security analytics is characterized by heterogeneous data, data warehouses, technologies, system monitoring, and dashboards [3]. It demands descriptive, diagnostic, detective, predictive, and prescriptive operations [15].

REQUIREMENT 3 (KNOWLEDGE). Security analytics enables knowledge generation, which is key for sharing [4,24], whereby digital twins enrich security analytics.

3.4. Intertwining Environments

We denote the primary entities and their relationships using an entity-relationship model (cf. Figure 3) in preparation for the formal model. Gray-colored blocks and dashed lines represent the logical components. Here, we represent the physical environment and the virtual environment as the main components in Industrial IoT. It should be noted that the lifecycle participants and the digital twin data can be mapped to both the physical and virtual environments. However, this plays only a minor role. Furthermore, we define security analytics as an internal component of the virtual environment [23]. The digital twin can be either in one of the three modes of operation: simulation, replication, or analytics.

The physical environment contains physical assets, i.e., IoT devices. A physical asset goes through its lifecycle with different lifecycle participants intertwined. A physical asset is assigned for a lifecycle phase at a given time, and other assets may pass through the same phase. Since digital twins are physical assets, they follow their physical counterparts through the lifecycle. They interact with the same and possibly other lifecycle participants that fit into the current lifecycle phase of the physical asset. In addition, digital twins have specific data, especially metadata, state, and historical data. Security analytics is placed above the digital twin, benefits from the semantics layer, and processes a set of digital twin data.

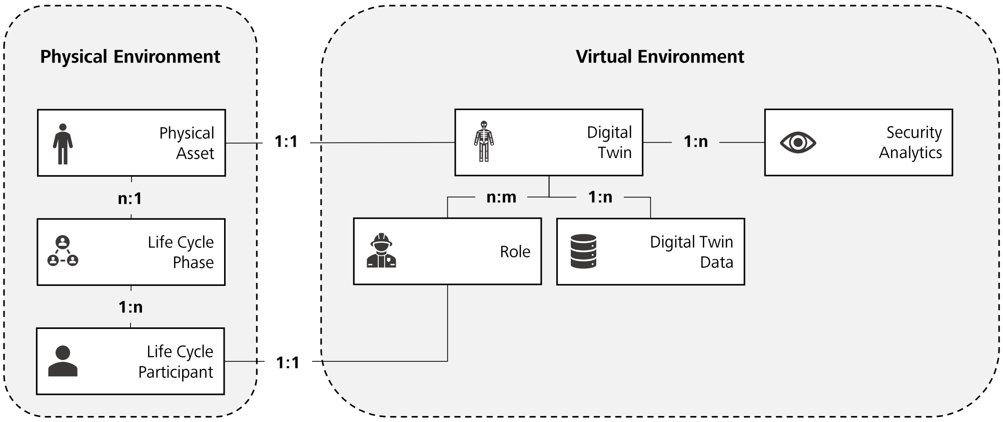


Figure 3. Security analytics through the intertwining of physical assets and digital twins.

4. Formal Model

We develop the formal model for digital-twin-based security analytics in the Industrial IoT in the following. This formal model is key to approaching the DT2SA model that supports digital-twin-based knowledge generation and sharing, enabling collaboration between lifecycle participants. Note that this model only abstractly illustrates which data can be coupled to which operation mode of the digital twins and security analytics operation. In Section 4.1, we define the formal model for the physical environment, and in Section 4.2, that for the virtual environment also containing security analytics described in Section 4.3.

4.1. Physical Environment

The physical environment comprises the physical asset and its associated lifecycle phases. We view physical assets as tangible and intangible artifacts of an organization. The term physical is used to distinguish between real-world and virtual-environment assets. It includes processes, software, and hardware (i.e., IoT). Thus, we define physical IoT assets as $I := \{i_1, ..., i_n\}$.

LIFECYCLE. These physical assets progress through different lifecycle phases—e.g., from design to operation to recycling. We define the lifecycle phases as $L := \{l_1, ..., l_m\}$ and assign exactly one lifecycle phase at a given time e to a physical asset: $f : i^e \mapsto l \mid i \in I \land l \in L$. Within a lifecycle phase, participants interact with physical assets, which we define as follows: $P := \{p_1, ..., p_k\}$. Due to this formal definition, a physical asset is always situated in exactly one phase of the lifecycle at a given time and interacts with the respective lifecycle participants.

4.2. Virtual Environment

The virtual environment describes the digital twin and the relation to several lifecycle participants and respective data. This paper considers the digital twin as a mapping to the physical asset. As the digital twin operates one of the three operation modes, namely, analytics, replication, and simulation [28], we define the operation mode analytics as T_{ana}, the replication as T_{rep}, and the simulation as T_{sim}. All digital twins are summarized through $T := \{t_1, ..., t_l\}$ and mapped to the physical assets as follows: $g : t \mapsto i \mid t \in T \land i \in I$. As digital twins T are mapped to physical assets A, and these to lifecycle phases L, we conclude $(f \circ g) \Leftrightarrow T \Rightarrow L$. Thus, the digital twin is in the same lifecycle phase as the physical asset.

DIGITAL TWIN DATA. We further define the digital twin data as $D := \{d_1, ..., d_p\}$. As already depicted, digital twins have several types of data, namely, descriptive and dynamic asset data, dynamic environment data, historical asset data, and semantics. Descriptive data D_{desc} includes specifications and static data of a physical asset, such as year of manufacture, model number, or identifiers. Dynamic asset data (we refer to as state data D_{state}) bundles all operational and asset-specific data that virtually represent assets given data structures dynamically. Dynamic environment data D_{env} are external stimuli that surround physical assets, e.g., data captured through external interfaces (i.e., sensors). Dynamic environment data do not only refer to data captured by the asset, e.g., network data. Historical asset data D_{hist} are a collection of persisted data. Semantics D_{sem} represent models that yield relevant data relations (this is also knowledge and wisdom, according to the DIKW). Note that it is not possible to have machines produce wisdom. Wisdom is also perceived differently by each lifecycle participant. Last, we define security-related data as D_{sec}. The latter holds all data for security operations, e.g., policies, SIEM rules, signatures, or threat intelligence data. These data types relate to each other. We define D_{hist} as the historical data storage that aggregates various types of data, whereby $\exists d \in D(d \in D_{state} \vee d \in D_{env} \vee d \in D_{sec})$. D_{desc} is considered not to change over time. The data types D_{state}, D_{env}, and D_{sec} are historically stored and a subset of data type D_{hist} is exceeded if a certain threshold s is breached (i.e., the arrival time of data d^t). We specify this exceeding or persistence as follows, where s is a timely restricted threshold:

$$persist(d_i) = \begin{cases} d_i \in (D_{state} \vee D_{env} \vee D_{sec}) & if\ d^t\ >\ s \\ d_i \in D_{hist} & otherwise. \end{cases}$$

4.3. Security Analytics

As described above, we consider security analytics as big data analytics from a cybersecurity perspective, distinguishing between analytical operations, namely, descriptive, diagnostic, detective, predictive, and prescriptive [3,15]. We define all analytical operations as $A := \{a_1, ..., a_q\}$, whereby $a \in A$ holds the respective results of one analytical operation (generated knowledge). Additionally, analytical operations O yield subsets, i.e., analytical operations categories: $(A_{desc} \cup A_{diag} \cup A_{det} \cup A_{pred} \cup A_{pres}) \subseteq A$. We consider all analytical categories, as we provide an approach to holistically integrating security analytics. Digital-twin-based security analytics involves knowledge and bidirectional communication with the physical environment.

KNOWLEDGE. We assume data, information, knowledge, and wisdom as the foundation of analytical operations but do not distinguish between these knowledge types as described by Böhm et al. [4]. We consider the sensemaking process a black box, as it is hard to define rationales, but we include cybersecurity knowledge. We refer to specific knowledge about observables as K_{obs}, knowledge about indicators of compromise as K_{ind}, and learning about incidents as K_{inc}.

INCIDENT RESPONSE. One of the core properties of digital twins is the bidirectional communication to real-world assets. Therefore, we also specify incident response. Incident response is part of security analytics and involves security orchestration and response. Thereby, playbooks are vital defining actions $C := \{c_1, ..., c_s\}$ given certain knowledge about incidents K_{inc} and threats. Integrating digital twins T into the incident response process is called orchestration, and a reaction to an event using digital twins is a response. A response to an incident is triggered by an event and solved through the orchestration of digital twins. Digital twins enable the orchestration, so we define the orchestration O concerning digital twins as $O \circ T$, and the response is as follows:

$$response := ((k \mapsto ot), c \mid c \in C, k \in K_{inc}, ot \in (O \circ T))$$

DESCRIPTIVE ANALYTICS. Descriptive analytics descriptively summarizes the data in context, using visualizations and statistical methods. This analytical operation relies on historical data to provide hindsight and identifies relevant observables. Descriptive analytics requires the operation mode T_{ana}. We define a descriptive operation as follows:

$$A_{desc} := (d, t \mid d \in D_{hist}, t \in T_{ana}) \mapsto K_{obs}$$

DIAGNOSTIC ANALYTICS. Diagnostic analytics goes beyond descriptive analytics and investigates the causes of phenomena (i.e., incidents). Due to this retro perspective, diagnostic operations incorporate historical data and identify indicators of compromise in observables. Diagnostic operations also require the operation mode T_{ana} and integrate knowledge about observables K_{obs}. We define a diagnostic operation as follows:

$$A_{diag} := (d, k, t \mid d \in D_{hist}, k \in K_{obs}, t \in T_{ana}) \mapsto K_{ind}$$

DETECTIVE ANALYTICS. Detective analytics generates new insights and relies on methods such as signatures or rules from a cybersecurity perspective. This analytical operation is often termed discovery analytics in big data analytics and links insights from historical data. We evolved this idea by appending insights from real-time data, which is highly relevant in the case of real-time event correlation (i.e., SIEM). Detective analytics can be in either operation mode, T_{rep} or T_{sim}. A detective operation links several indicators of compromise to incidents. It involves knowledge about observables to generate knowledge about incidents K_{inc}. This knowledge defines analytical and SIEM typical measures, e.g., signatures or rules. Detective operations rely on all kinds of data. We define detective operations as follows:

$$A_{det} := (k, t \mid k \in K_{ind}, t \in T_{rep} \vee t \in T_{sim}) \mapsto K_{inc}$$

PREDICTIVE ANALYTICS. Predictive analytics utilizes data and knowledge to predict the future. Thereby, predictive operations deploy semantics (i.e., mathematical models or simulations) and involve knowledge about incidents to predict if an incident is likely to happen. Predictive operations rely on the results of detective operations A_{det} and create new incident knowledge K'_{inc}. Predictive operations can also be in operation mode T_{rep} or T_{sim}. Predictive operations also rely on all kinds of data. We define a predictive operation as follows:

$$A_{pred} := (a, k, t \mid a \in A_{det}, k \in K_{inc}, t \in (T_{rep} \oplus T_{sim})) \mapsto K'_{inc}$$

PRESCRIPTIVE ANALYTICS. Prescriptive analytics identifies, evaluates, and suggests appropriate security orchestration to mitigate an incident. Thereby, simulations play a decisive role in deriving decisions from the different scenarios, e.g., through what–if simulations. Prescriptive operations can also be either in operation mode T_{rep} or T_{sim} and require results of the analytical operations A_{det} or A_{pred}. Further, prescriptive operations identify appropriate actions for incident response activities out of existing knowledge about incidents K_{inc}. Prescriptive operations may encompass all kinds of data. We define a prescriptive operation as follows:

$$A_{pres} := (a, k, t \mid a \in (A_{det} \oplus A_{pred}), k \in K_{inc}, t \in (T_{rep} \vee \in T_{sim})) \mapsto C$$

KNOWLEDGE SHARING. We adopt parts of the data-sharing concept from Dietz et al. [29] in our sharing principles for digital twin data. We add the sharing of cybersecurity knowledge K and security-related data D_{sec}. We assume that roles have access to digital twins, dependent on permissions. If permission is granted, roles are invited to access digital twin data, including existing knowledge and models. Further, lifecycle participants can contribute expertise and write relevant descriptive and security-related data entries or semantics.

$$read := (t, d, k \mid t \in T, d \in D, k \in K)$$

$$write := (t, d, k \mid t \in T, d \in (D_{desc} \oplus D_{sem} \oplus D_{sec}), k \in K)$$

5. DT2SA Model

Based on the formal model in Section 4, we envision the DT2SA model. Figure 4 shows the respective model. We adopt the functional view of the Industrial Internet Reference Architecture (IIRA) [30], based on the five domains: the control domain, information domain, operations domain, functional domain, and business domain. This reference architecture ensures the DT2SA model is embedded in the IoT. In the following, we explain the respective domains of our DT2SA model.

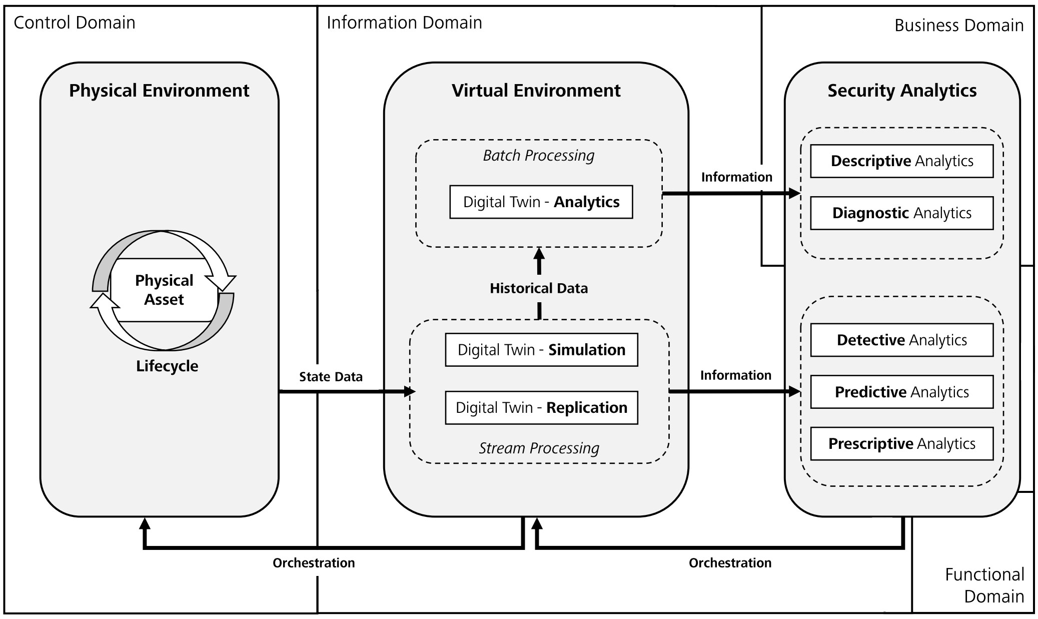


Figure 4. DT2SA model.

Control Domain. The control domain seamlessly captures physical assets and their data. Physical assets refer to artifacts in the Industrial IoT, such as processes, machines, or plants. In this context, physical assets collect data through essential interactions, such as machine-to-machine communication. Asset-specific data of interest includes descriptive, state-related, and security-related data. Physical assets, i.e., machines, are arranged in networks. These networks also generate operational and security-related data, such as network traffic. However, the overarching task of the physical environment is to make all relevant data from the physical assets and their environment available to their respective digital twins. The control domain uses various networks for data transmission. The data pass through so-called proximity, access, and service networks. Proximity networks are responsible for short-range communication and access networks for long-range communication. Service networks are enterprise networks that handle communications between business applications.

Information Domain. The information domain is responsible for data acquisition, processing, and persistence. The physical environment feeds the respective digital twins with data and receives data and commands from the digital twins. Digital twins thus claim bidirectional communication with the physical environment. Technically, bidirectional communication is realized by so-called event-based messaging platforms that rely on hubs or brokers based on publish/subscribe messaging. However, the information domain includes historical, state, and semantic data. The use of different data stores manages the diversity of data. The digital twin in analytics mode processes historical data in a batch-processing pipeline. Digital twins using replication and simulation modes rely heavily on real-time data in a stream processing pipeline. Each pipeline supports different analytical operations. Descriptive and diagnostic operations are coupled with the digital twin's analytics mode to provide hindsight. Detective operations are time-dependent and should be coupled with the replication mode. The digital twin constantly feeds detective operations with real-time state data. Predictive and prescriptive operations build on both simulation and replication modes. Both modes of operation are the foundation for predicting incidents or recommending incident response operations.

FUNCTIONAL DOMAIN. The functional domain operates through rule-based decisions. From a cybersecurity perspective, we map incident response activities to the functional domain because incident response orchestrates assets using predefined logic. For example, incident response is automated using playbooks to secure the IoT [12]. Digital twins provide an additional layer of security and orchestrate physical assets. Incident response is also event-based, receiving information and knowledge from human experts and digital-twin-based security analytics. However, the incident response process is triggered by analytical operations.

BUSINESS DOMAIN. The business domain includes descriptive and diagnostic operations. The business domain promotes the exchange and generation of knowledge. For example, visualized information obtained from descriptive and diagnostic operations helps to identify correlations and gain insights. This knowledge can further optimize digital twin models and analytical operations. We see the potential of a dedicated knowledge-sharing and management platform to explore previously untapped knowledge in cybersecurity visually.

6. Proof of Concept

We validate our DT2SA model in regard to its applicability to an ongoing research project. In Section 6.1, we investigate to what extent all requirements of the DT2SA model have been met. Next, we validate our prototype TWINSIGHT in Section 6.2 and evaluate the applicability of our model to the existing literature on digital-twin-based security analytics in Section 6.3. In Secion 6.4, we define our experimental setting leading to results in Section 6.5, which we discuss thoroughly in Section 6.6.

6.1. DT2SA *Components*

In Section 3.3, we defined the key requirements for the DT2SA model. In the following, we assess and discuss the fulfillment of each requirement.

DIGITAL TWIN (R1). The DT2SA model includes physical assets, virtual representations, communication management, data management, and modes of operation. It seamlessly integrates physical assets into digital twins, making virtual representations indistinguishable from physical assets through synchronization and communication. It also addresses all relevant mechanisms for data management, including how to retain and process data. All requirements for the digital twin are met.

SECURITY ANALYTICS (R2). The DT2SA model incorporates security analytics through the variety of data and semantics used by analytical operations. The model also incorporates different data from different data stores. Although technologies are essential to the implementation of the model, less attention has been paid to them. We also touched on the incident response process that manifests the interaction with the physical asset and the orchestration of cybersecurity operations. We considered all analytical operations to cover security analytics fully. Less attention was paid to interactive dashboards and system monitoring, as these functional components can be effortlessly added and only make the model unnecessarily overcomplex. However, our model covers almost all features of security analytics.

KNOWLEDGE (R3). The DT2SA model enables the generation and sharing of knowledge among lifecycle participants. For cybersecurity knowledge generation and sharing, the model proposes the integration of key lifecycle participants through digital twins and security analytics. The model also embeds the knowledge hierarchy through the formal model. Information sharing is not addressed in detail.

The DT2SA model addresses almost all requirements and provides relevant perspectives on digital twins and security analytics. Although the requirements are met, there is a lack of applicability in practice. Therefore, we refer to an ongoing research project and implement TWINSIGHT to validate our model further.

6.2. Use Case: SISSeC

To further validate the DT2SA model, we implement a microservice architecture called TWINSIGHT. TWINSIGHT relates to SISSeC, an ongoing research project in Germany. SISSeC aims to securely connect a printed circuit board (PCB) manufacturer's machine and sensor data to a cloud via an edge gateway. The PCB manufacturer's overall goal is to collect data and predict likely future operating conditions of machines with digital twins. From a cybersecurity perspective, these digital twins should enable intrusion detection. The project also aims to make security-relevant data available to lifecycle participants via a marketplace. In this way, lifecycle participants and third parties can share their knowledge.

We instantiate the DT2SA model for SISSeC. The PCB manufacturer requires state data, descriptive data, and environmental data to use our model. We focus on a drilling and milling machine and its system-specific operational data. This machine sends data over the proximity network to the edge node, which forwards the data to the cloud. The cloud contains digital twins that enable bidirectional communication, and thus, control of the machines. To instantiate the DT2SA model, we formulate two main objectives of TWINSIGHT: incident detection and threat detection. Then, we choose the analytical operation that satisfies this goal: the detective operation. The next step is to choose an appropriate operation mode for the digital twin.

6.3. Applicability

Before specifying a concrete operation mode for detective operations in SISSeC, we validated the applicability of our DT2SA model to existing literature dealing with digital twins, in particular, intrusion or anomaly detection. Given our use case, we only considered operational modes that enable detective operations, so we focused on simulation and replication modes. In the following, we list relevant literature that presents concrete instantiations of our DT2SA model:

- *Simulation mode* [9,22,23,31–39];
- *Replication mode* [21,36,40–42].

We found that more papers deal with simulation mode and less deal with replication mode. This literature can also serve as references to provide concrete ideas on using simulation- or replication-based digital twins for security analytics. For example, simulation-based digital twins are utilized for incident prediction [35], and state-replication-based digital twins are used for intrusion detection [21]. These approaches fit into the DT2SA model and follow the same scheme: digital twins organize data and models and make them available for subsequent analytic operations. In simulation models for machine learning training, we found, among other things, historical data. Most of the data used are real-time or specification data. Based on the SISSeC use case, the fact that replication-based digital twins for IDS have not been explored, and the lack of sharing capabilities, we develop an experimental environment for replication-based digital twins for detective operations.

6.4. Experimental Setup

We implement TWINSIGHT for digital-twin-based incident and threat detection containing detective operations. TWINSIGHT is publicly available on GitHub. Figure 5 shows the settings in TWINSIGHT. The developments in TWINSIGHT are driven by the SISSeC use case and are intended to illustrate how detective security analytics can be implemented with replication-based digital twins. Note that this experimental setup focuses on only a specific part of the DT2SA framework: replication-based digital twins and detective operations. In this way, we aim to gain new insights into replication-based digital twins and detective security analytics and determine whether the DT2SA model covers these aspects. In the following, we describe these components in more detail.

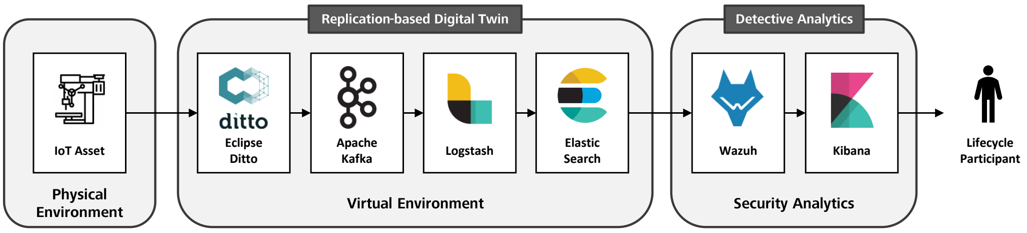


Figure 5. TWINSIGHT setting implemented in the research project SISSeC.

PHYSICAL ENVIRONMENT. The physical environment consists of identical drilling and milling machines relying on Raspberry Pis for data transfer. Depending on the setup, a machine consists of several Raspberry Pis. We use Raspberry Pi 3B+ models running Raspbian GNU/Linux 11 with 1GB RAM. These devices are interconnected using an event-based messaging architecture. More specifically, the devices communicate via MQTT 3.1 with an MQTT broker. In doing so, each client implements the Paho MQTT client library in Python. We can now use MQTT to collect device-related data, such as the state of an asset. Interesting setup details follow in the virtual environment.

VIRTUAL ENVIRONMENT. Eclipse Ditto is the digital twin software built on an MQTT broker that manages messages received from clients. Eclipse Ditto implements an event-based API that allows the definition of device representations and messages. We model each drilling and milling machine in Eclipse Ditto using the built-in JSON schema. In addition, Eclipse Ditto allows us to define connectors, which we use as a bridge to the MQTT broker. Of course, messaging would allow for device orchestration and response, but we focus only on security analytics, not incident response. As Eclipse Ditto only stores the current state of a machine, Apache Kafka collects real-time data; and Logstash subscribes to all topics, transforms messages, and stores them in Elasticsearch. These applications are deployed on a virtual machine running Ubuntu 20.04.3 LTS with 12 GB of RAM, six cores, and 60 GB of storage.

SECURITY ANALYTICS. We implement detective security analytics using Wazuh and its native Kibana integration. Wazuh is open-source software that performs sophisticated security operations and incident response. It relies on agents installed on the assets to be monitored and orchestrated. Kibana is used only for the virtual representation of agent-based information. Both applications also run on the same machine as the virtual environment.

6.5. Results

In implementing replication-based digital twins for detective security analytics, we gained insights and results that we would like to share. Our research provides results related to implementation, security analytics, digital twins, the IoT, and knowledge sharing.

IMPLEMENTATION. We support lifecycle participants by aligning digital twins and their modes of operation with security analytics. The formal model makes implementing software based on digital twins more feasible. We have found that replication-based digital twins fit real-time data processing, and the literature confirms that simulation works decoupled from the live system, e.g., Dietz et al. [43]. Using an event-based microservice architecture ensures flexibility and real-time data processing. We have also found that selecting an appropriate mode of operation, assets, and data is more efficient when starting with the desired goal of security analytics. We consider this to be top-down DT2SA.

SECURITY ANALYTICS. Since most requirements are sufficiently met, the DT2SA model does not lack essential components. In particular, more security monitoring and an interactive dashboard can easily be added. Nevertheless, these features are not considered core features of security analytics. We also found that security analytics technologies, e.g., Wazuh, do not work properly with device representations. They are designed to access a machine's resources using agents. In modern organizations, machines no longer consist of a single component but form complex systems of systems. In such environments, these components should be monitored in relation using security analytics. Appendix A shows the Wazuh user interface and two views that visualize the detection of incidents and threats in real-time and point out the problem at hand. Figure A1 shows all integrated assets and threats with Wazuh agents. Figure A2 shows all attacks against a specific asset. Looking at complex systems that consist of multiple components (and agents) makes traditional security analytics inefficient. Security analytics should integrate asset representations and exploit relationships between assets when considering digital twins. However, correlating events related to complex systems of systems and digital twins is paramount. More efficient analytics should enable the mapping and correlation of agents from assets to digital twins to realize their full potential. Nonetheless, security analytics would benefit from systems of systems approaches for improving the visualization and resolution of security events.

DIGITAL TWINS. Digital twin software allows for easier data processing and analysis, as states are updated dynamically. Eclipse Ditto, a replication-based digital twin focusing on states and functions, provides dynamic user management that defines roles and their privileges.

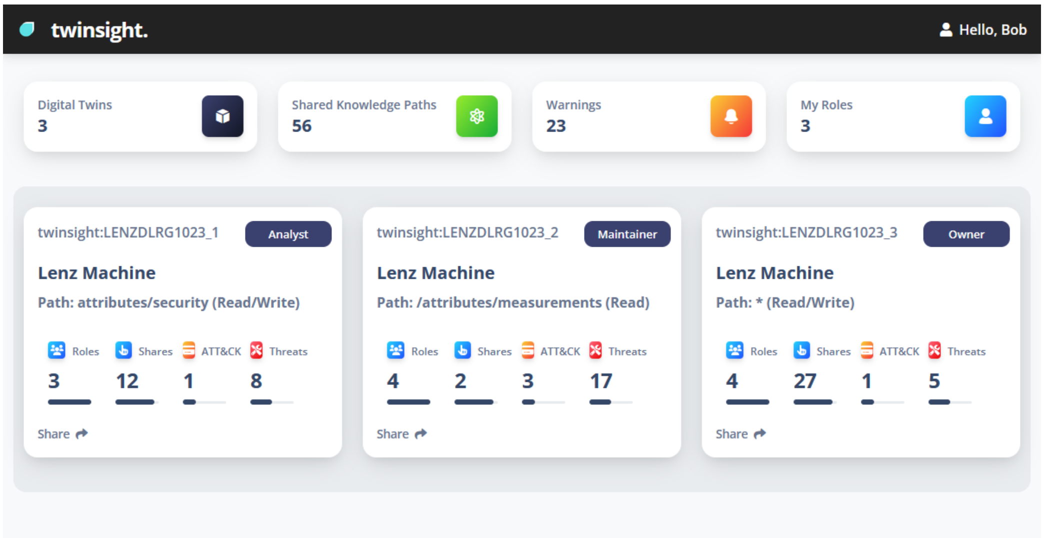
We found that built-in user management enables fine-grained sharing of digital twin data among lifecycle participants. We also found a research gap on digital twins in analytics mode. In addition, we learned in SISSeC that it is possible to implement certain application scenarios with digital twin software. Nevertheless, security analytics is only one of many possible application scenarios for digital twins. Thus, TWINSIGHT can be extended to other analytic operations, and even to more sophisticated security operations.

In addition, digital twins are ideally suited as the ground truth of knowledge. We consider digital twins as hubs for managing device-specific and security-related data. As discussed earlier, they ensure the correct mapping of components, which allows us to model system of systems. Listing 1 shows an Eclipse Ditto device representation, including security-related information and analytics results. In our use case, we included relevant common platform enumerations (CPE) to query vulnerabilities related to the digital twins' components. We can also map IDS agents to a specific digital twin to include the most recent alert in the device representation. The application scenarios and extensions of security analytics are numerous and not limited to component mapping, vulnerability queries, or relationships with IDS agents.

Listing 1. Digital Twin Definition in Eclipse Ditto.

```
{
"thingId": "SISSeC:Lenz_DRB610_1",
"policyId": "SISSeC:policy",
"attributes": {
"manufacturerID": "4302",
"manufacturerName": "Manufacturer",
"dateCode": "20160516",
"model": "Model",
"type": "Drill & Mill Machine",
"image": "/resources/...",
"location": "Hall 1",
"measurements": {...},
"security": {
"cpe": [
{
"name": "Raspberry Pi Model 4",
"usage": "Edge Device",
"enum": "cpe:2.3:h:raspberrypi:raspberry_pi_4_model_b:-:*:*:*:*:*:*:*"
}
],
"alert": {
"sensor": "A002",
"msg": "test alert",
"src_ip_mac": "10.10.10.10",
"dst_ip_mac": "10.10.10.20",
"src_port": "123",
"dst_port": "5065",
"time": "27/08/2022",
"packet_len": "80",
"protocol": "TCP",
"ether_type": "0000"
}
}
}
}
```

Digital twins take security analytics to a new level in knowledge generation and sharing. As shown in Figure 6, our TWINSIGHT UI leverages security-related knowledge from digital twins to inform security professionals and enable knowledge sharing. The user interface integrates device representation (digital twin) and highlights current threats, potential attacks, and granted shares. We believe system-of-systems alerts would lead to more intuitive interaction with Wazuh security analytics software by reducing complexity through digital twins. The alerts would point directly to a system-of-systems component and help understand the big picture and impact of threats and incidents.



twinsight.

Hello, Bob

Digital Twins 3

Shared Knowledge Paths 56

Warnings 23

My Roles 3

twinsight:LENZDLRG1023_1 Analyst
Lenz Machine
Path: attributes/security (Read/Write)
Roles 3 | Shares 12 | ATT&CK 1 | Threats 8
Share

twinsight:LENZDLRG1023_2 Maintainer
Lenz Machine
Path: /attributes/measurements (Read)
Roles 4 | Shares 2 | ATT&CK 3 | Threats 17
Share

twinsight:LENZDLRG1023_3 Owner
Lenz Machine
Path: * (Read/Write)
Roles 4 | Shares 27 | ATT&CK 1 | Threats 5
Share

Figure 6. TWINSIGHT UI combining security analytics and digital twins.

INTERNET OF THINGS. Indeed, Wazuh relies heavily on an agent-based architecture. Due to its dependence on agents, not all IoT devices are sufficiently addressable. In IoT, a distinction is made between controllable and addressable devices. Addressable devices are reachable via IP addresses and controllable via a controller, i.e., a hub. Wazuh only allows observing addressable devices and leaves out controllable devices, so IoT security analytics still differs from traditional approaches. Controllable devices must be monitored differently to detect and mitigate threats or incidents efficiently.

KNOWLEDGE SHARING. Knowledge sharing in cybersecurity is still in its infancy but will likely increase in the coming years. For example, we found it still difficult to obtain data from machines or their components because the interfaces are not standardized or intentionally hidden. We encountered this problem when we tried to access the interfaces of the machines to get the state data for the digital twin. It took a few months before we could get initial access and start modeling the digital twins. We received excellent feedback from the SISSeC working group, particularly on digital twins as a cybersecurity knowledge generation and sharing facilitator. In this context, data will be shared with the machine maintainer. The PCB manufacturer plans to use parts of TWINSIGHT to generate and share knowledge between the machine owner and IT maintenance. This means that problems can be addressed remotely and do not necessarily have to be solved on-site.

6.6. Discussion

Our research contributes to the scientific community and industry. We also point out the limitations of our research.

SCIENTIFIC CONTRIBUTION. We have summarized research on big data analytics and linked knowledge generation and sharing to cybersecurity. We have formulated a model that envisions digital twins for more efficient security analytics in organizations. We have demonstrated the workings of the digital twin and clarified that simulations are not the only contributors to cybersecurity operations. The overall model helps researchers understand digital twins and aims to draw attention to the use of digital twins for security operations, especially analytics.

PRACTICAL CONTRIBUTION. We have implemented a microservice architecture demonstrating replication-based digital twins for security analytics. We have shown how threat and incident detection can be handled using digital twins, providing a playground for further use cases. Organizations can leverage open-source technologies to deploy their digital twins using the DT2SA model or build sophisticated use cases from scratch using our model. We also highlighted the digital twin paradigm to increase user adoption and make the IoT even more secure. We also drew attention to security analytics technologies, e.g., Wazuh, and their lack of abstraction to form even more complex system of systems. These technologies should take steps toward an asset-specific view that allows users to define complex systems of systems connected to Wazuh agents.

LIMITATIONS. During our research, we encountered several challenges that needed to be solved. Due to the different solution strategies and limited resources, there are limitations to our research. We did not elaborate on access control models. These models provide a clear perspective on inherent roles and grant data access to digital twins. In addition, we validated our model only concerning detective operations. We can only estimate the feasibility of other analytical operations and refer to other literature. However, further research is needed to validate our model in more detail.

7. Conclusions and Future Work

This research aims to leverage cybersecurity knowledge to secure the IoT. We promote cybersecurity knowledge generation and sharing by aligning security analytics with digital twins. Digital twins enable security analytics with high fidelity because they bring semantics and exploit bidirectional communication with their physical counterparts. They take security analytics to a new level, enabling lifecycle centrality and integration among lifecycle participants. This integration promotes the secure sharing of cybersecurity knowledge, such as security states, misconfigurations, or vulnerabilities.

We answered the research question *"How can one align security analytics and digital twins?"* by starting with the foundations of knowledge generation and sharing. We then defined a formal model that elaborates the DT2SA model for adapting security analytics to digital twins. To our knowledge, the DT2SA model is the first to define security analytics comprehensively. We contributed to best practices for research and organizations and bridged the gap between them. Our open-source microservice architecture TWINSIGHT demonstrated practical feasibility and is a starting point for on-building analytical operations. We want to highlight possible future research directions:

- Future research should address decision support for selecting digital twin modes and analytic operations. In particular, whether an analytic operation supports a particular application scenario should be investigated. The goal is to assist analysts in selecting appropriate operation modes for their scenarios. However, the digital twin offers significant cybersecurity opportunities that need to be more fully explored and exploited.
- There is still a considerable need for research, especially in the area of security analytics, since research has focused only on intrusion detection. For example, research should address different analytics implementations based on digital twins. In particular, security monitoring for IoT is urgently needed, as heterogeneous IoT assets form opaque IoT networks. In addition, security analytics research should compare traditional security analytics approaches, such as those implemented in Wazuh, with system-of-systems approaches. It is of the highest interest to evaluate whether analysts using system-of-systems approaches are even more efficient at detecting incidents. Our TWINSIGHT UI highlights opportunities for this evaluation. In addition, there is a significant need for research in implementing a Wazuh plugin for modeling complex system of systems. Finally, future research should work to leverage digital twin recommendations to secure controllable and addressable IoT networks proactively.

While security analytics generates knowledge, digital twins improve overall knowledge generation and enable cybersecurity knowledge sharing. Organizations should leverage cybersecurity knowledge and focus more on digital twins and security analytics. In addition, supply chains should pay more attention to digital twins and their potential for cybersecurity to address sophisticated attacks and APTs. We believe that the digital twin (system of systems) will continue to emerge as a cornerstone of collaboration between lifecycle participants by leveraging cybersecurity knowledge in the Industrial IoT.

Author Contributions: P.E. carried out DT2SA development and the definition of the formal model. P.E. also contributed to knowledge generation and sharing in the cybersecurity domain. G.P. participated in the definition of the DT2SA model. P.E. implemented the microservice architecture TwinSIGHT. P.E. and G.P. also helped to draft the manuscript and revised it critically for important intellectual content. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) and the Bavarian Ministry of Economic Affairs, Regional Development and Energy (BayStMWi). Funding numbers: 16KN085725 and DIK0338/01.

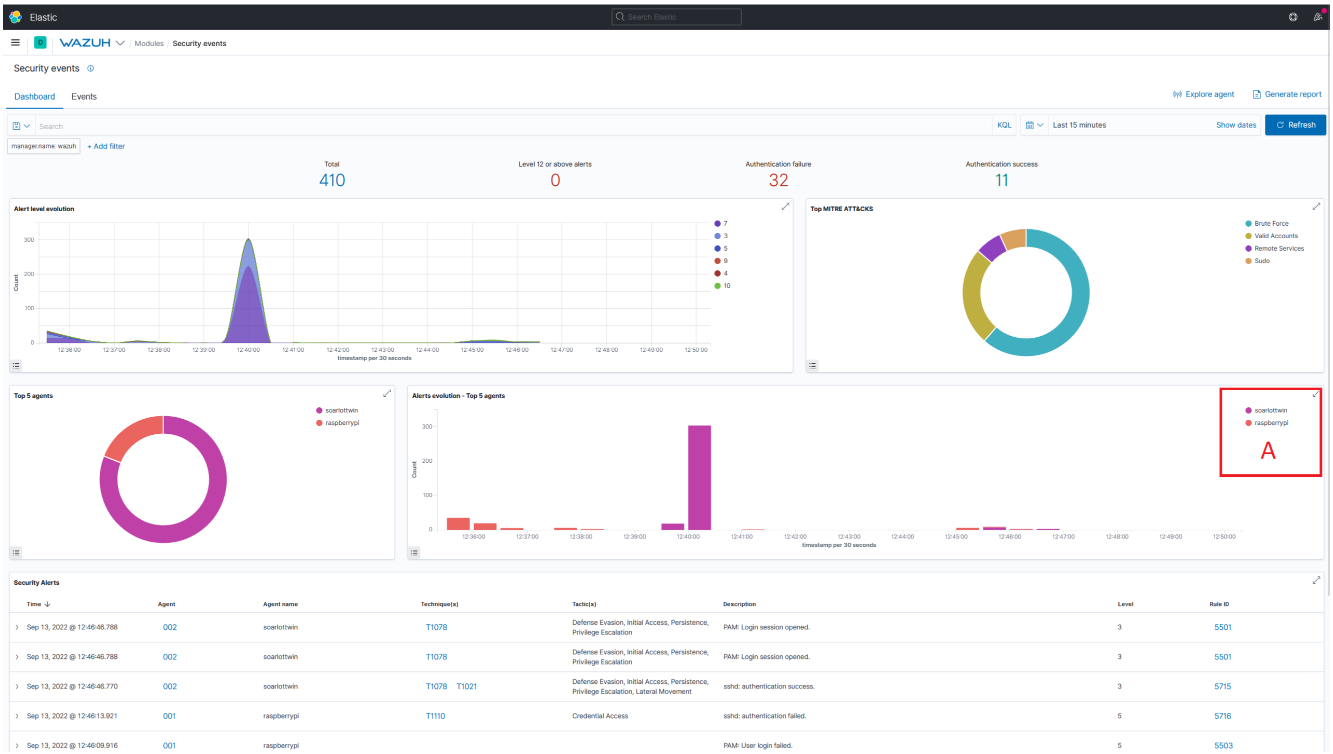
Data Availability Statement: Our data is available on GitHub: https://github.com/philipempl/twinsight.

Acknowledgments: We want to acknowledge all project partners involved in the Secure Industrial Semantic Sensor Cloud (SISSeC) project funded by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) as part of its central innovation program for small and medium-sized businesses. This research was also partly supported by the Bavarian Ministry of Economic Affairs, Regional Development and Energy (BayStMWi), as part of the INSIST project.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Security Analytics Using Wazuh

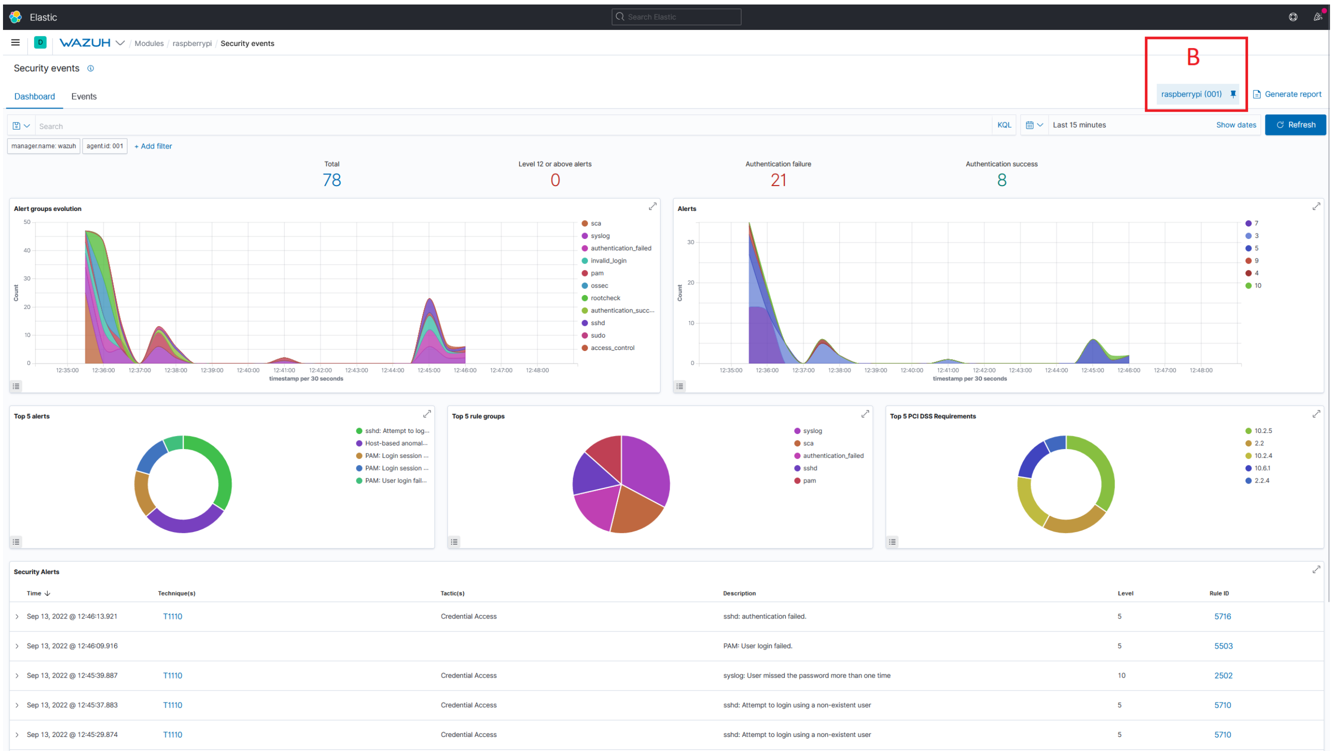
This appendix shows the need for integration and modeling a complex system of systems in Wazuh. Since machines nowadays consist of multiple components, security analytics technology should provide semantic modeling capabilities to analyze systems of systems and their respective components.



Security Alerts

Time	Agent	Agent name	Technique(s)	Tactic(s)	Description	Level	Rule ID
Sep 13, 2022 @ 12:46:46.788	002	soarlottwin	T1078	Defense Evasion, Initial Access, Persistence, Privilege Escalation	PAM: Login session opened.	3	5501
Sep 13, 2022 @ 12:46:46.788	002	soarlottwin	T1078	Defense Evasion, Initial Access, Persistence, Privilege Escalation	PAM: Login session opened.	3	5501
Sep 13, 2022 @ 12:46:46.770	002	soarlottwin	T1078 T1021	Defense Evasion, Initial Access, Persistence, Privilege Escalation, Lateral Movement	sshd: authentication success.	3	5715
Sep 13, 2022 @ 12:46:13.921	001	raspberrypi	T1110	Credential Access	sshd: authentication failed.	5	5716
Sep 13, 2022 @ 12:46:09.916	001	raspberrypi			PAM: User login failed.	5	5503

Figure A1. Analyzing all events of all assets in Wazuh (A shows all agents).



raspberrypi (001)

manager.name: wazuh | agent.id: 001

Last 15 minutes

Total 78 | Level 12 or above alerts 0 | Authentication failure 21 | Authentication success 8

Security Alerts

Time ↓	Technique(s)	Tactic(s)	Description	Level	Rule ID
Sep 13, 2022 @ 12:46:13.921	T1110	Credential Access	sshd: authentication failed.	5	5716
Sep 13, 2022 @ 12:46:09.916			PAM: User login failed.	5	5503
Sep 13, 2022 @ 12:45:39.887	T1110	Credential Access	syslog: User missed the password more than one time	10	2502
Sep 13, 2022 @ 12:45:37.883	T1110	Credential Access	sshd: Attempt to login using a non-existent user	5	5710
Sep 13, 2022 @ 12:45:29.874	T1110	Credential Access	sshd: Attempt to login using a non-existent user	5	5710

Figure A2. Analyzing events of one asset in Wazuh (B showcases the agent filter).

References

1. ENISA. *Threat Landscape for Supply Chain Attacks*; Technical report; ENISA: Athens, Greece, 2021.
2. Ardagna, C.; Corbiaux, S.; Sfakianakis, A.; Douligeris, C. *ENISA Threat Landscape*; Technical report; ENISA: Athens, Greece, 2021.
3. Mahmood, T.; Afzal, U. Security Analytics: Big Data Analytics for Cybersecurity: A Review of Trends, Techniques and Tools. In Proceedings of the 2nd National Conference on Information Assurance (NCIA 2013), Rawalpindi, Pakistan, 11–12 December 2013; pp. 129–134.
4. Böhm, F.; Vielberth, M.; Pernul, G. Bridging Knowledge Gaps in Security Analytics. In Proceedings of the Proceedings of the 7th International Conference on Information Systems Security and Privacy, ICISSP 2021, Online Streaming, 11–13 February 2021; Mori, P., Lenzini, G., Furnell, S., Eds.; SCITEPRESS: Setubal, Portugal, 2021; pp. 98–108. [CrossRef]
5. Skouloudi, C.; Malatras, A.; Naydenov, R.; Dede, G. *Guidelines for Securing the Internet of Things*; Technical report; European Union Agency for Cybersecurity: Athens, Greece, 2020.
6. Pipikaite, A.; Bueermann, G.; Joshi, A.; Jurgen, J.; Bissell, K.; Aguirre, C.; Browder, T.; Pruitt, J. *Global Cybersecurity Outlook 2022: Insight Report*; Technical report; European Union Agency for Cybersecurity: Athens, Greece, 2022.
7. Boschert, S.; Heinrich, C.; Rosen, R. Next Generation Digital Twin. In Proceedings of the 12th International Symposium on Tools and Methods of Competitive Engineering (TMCE), Las Palmas de Gran Canaria, Spain, 7–11 May 2018; Horvath, I., Suarez Riviero, J., Hernandez Castellano, P., Eds.; TMCE 2020 Repository: Dublin, Ireleand, 2018; Volume 2018, pp. 209–218.
8. Eckhart, M.; Ekelhart, A. Digital Twins for Cyber-Physical Systems Security: State of the Art and Outlook. In *Security and Quality in Cyber-Physical Systems Engineering, With Forewords by Robert M. Lee and Tom Gilb*; Biffl, S., Eckhart, M., Lüder, A., Weippl, E.R., Eds.; Springer: Cham, Switzerland, 2019; pp. 383–412. [CrossRef]
9. Pokhrel, A.; Katta, V.; Colomo-Palacios, R. Digital Twin for Cybersecurity Incident Prediction: A Multivocal Literature Review. In Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops, 2020, ICSEW'20, Seoul, Republic of Korea, 27 June–19 July 2020; pp. 671–678. [CrossRef]
10. O'Connor, L. Strengthening Security with Digital Cyber Twins. 2021. Available online: https://www.accenture.com/us-en/blogs/technology-innovation/lisa-oconnor-strengthening-security-with-digital-cyber-twins (accessed on 29 May 2022).
11. Barricelli, B.R.; Casiraghi, E.; Fogli, D. A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. *IEEE Access* **2019**, *7*, 167653–167671. [CrossRef]
12. Empl, P.; Schlette, D.; Zupfer, D.; Pernul, G. SOAR4IoT: Securing IoT Assets with Digital Twins. In Proceedings of the 17th International Conference on Availability, Reliability and Security (ARES 2022), Vienna, Austria, 23–26 August 2022; Association for Computing Machinery: New York, NY, USA, 2022. [CrossRef]
13. Alcaraz, C.; Lopez, J. Digital Twin: A Comprehensive Survey of Security Threats. *IEEE Commun. Surv. Tutor.* **2022**, *24*, 1475–1503. [CrossRef]

14. Win, T.Y.; Tianfield, H.; Mair, Q. Big Data Based Security Analytics for Protecting Virtualized Infrastructures in Cloud Computing. *IEEE Trans. Big Data* **2018**, *4*, 11–25. [CrossRef]
15. Siow, E.; Tiropanis, T.; Hall, W. Analytics for the Internet of Things: A Survey. *ACM Comput. Surv.* **2018**, *51*, 74:1–74:36. [CrossRef]
16. Cárdenas, A.A.; Manadhata, P.K.; Rajan, S.P. Big Data Analytics for Security. *IEEE Secur. Priv.* **2013**, *11*, 74–76. [CrossRef]
17. Alguliyev, R.; Imamverdiyev, Y. Big Data: Big Promises for Information Security. In Proceedings of the 8th IEEE International Conference on Application of Information and Communication Technologies (AICT), Astana, Kazakhstan, 15–17 October 2014; pp. 1–4. [CrossRef]
18. Ackoff, R.L. From Data to Wisdom. *J. Appl. Syst. Anal.* **1989**, *16*, 3–9.
19. Empl, P.; Pernul, G. A Flexible Security Analytics Service for the Industrial IoT. In Proceedings of the 2021 ACM Workshop on Secure and Trustworthy Cyber-Physical Systems, Virtual Event, Charlotte, NC, USA, 28 April 2021; Gupta, M., Abdelsalam, M., Mittal, S., Eds.; ACM: New York, NY, USA, 2021; pp. 23–32. [CrossRef]
20. Menges, F.; Pernul, G. A comparative analysis of incident reporting formats. *Comput. Secur.* **2018**, *73*, 87–101. [CrossRef]
21. Eckhart, M.; Ekelhart, A. A Specification-based State Replication Approach for Digital Twins. In Proceedings of the 2018 Workshop on Cyber-Physical Systems Security and PrivaCy, CPS-SPC@CCS 2018, Toronto, ON, Canada, 19 October 2018; Lie, D., Mannan, M., Eds.; ACM: New York, NY, USA, 2018; pp. 36–47. [CrossRef]
22. Dietz, M.; Vielberth, M.; Pernul, G. Integrating Digital Twin Security Simulations in the Security Operations Center. In Proceedings of the 15th International Conference on Availability, Reliability and Security, Dublin, Ireland, 25–28 August 2020; Volkamer, M., Wressnegger, C., Eds., 2020, ARES '20, pp. 18:1–18:9. [CrossRef]
23. Damjanovic-Behrendt, V. A Digital Twin Architecture for Security, Privacy and Safety. *ERCIM News* **2018**, *2018*, 25–26.
24. Sacha, D.; Stoffel, A.; Stoffel, F.; Kwon, B.C.; Ellis, G.P.; Keim, D.A. Knowledge Generation Model for Visual Analytics. *IEEE Trans. Vis. Comput. Graph.* **2014**, *20*, 1604–1613. [CrossRef] [PubMed]
25. Preut, A.; Kopka, J.P.; Clausen, U. Digital Twins for the Circular Economy. *Sustainability* **2021**, *13*, 467. [CrossRef]
26. Putz, B.; Dietz, M.; Empl, P.; Pernul, G. EtherTwin: Blockchain-based Secure Digital Twin Information Management. *Inf. Process. Manag.* **2021**, *58*, 102425. [CrossRef]
27. Kritzinger, W.; Karner, M.; Traar, G.; Henjes, J.; Sihn, W. Digital Twin in Manufacturing: A Categorical Literature Review and Classification. *IFAC-PapersOnLine* **2018**, *51*, 1016–1022.
28. Dietz, M.; Pernul, G. Unleashing the Digital Twin's Potential for ICS Security. *IEEE Secur. Priv.* **2020**, *18*, 20–27. [CrossRef]
29. Dietz, M.; Putz, B.; Pernul, G. A Distributed Ledger Approach to Digital Twin Secure Data Sharing. In Proceedings of the Data and Applications Security and Privacy XXXIII—33rd Annual IFIP WG 11.3 Conference, DBSec 2019, Charleston, SC, USA, 15–17 July 2019; Lecture Notes in Computer Science; Foley, S.N., Ed.; Springer: Cham, Switzerland, 2019; Volume 11559, pp. 281–300. [CrossRef]
30. Lin, S.W.; Miller, B.; Durand, J.; Joshi, R.; Didier, P.; Chigani, A.; Torenbeek, R.; Duggal, D.; Martin, R.; Bleakley, G. *Industrial Internet Reference Architecture*; Technical report; Industry IoT Consortium: Boston, MA, USA, 2015.
31. Akbarian, F.; Fitzgerald, E.; Kihl, M. Intrusion Detection in Digital Twins for Industrial Control Systems. In Proceedings of the 2020 International Conference on Software, Telecommunications and Computer Networks (SoftCOM), Split, Croatia, 17–19 September 2020; pp. 1–6. [CrossRef]
32. Atalay, M.; Angin, P. A Digital Twins Approach to Smart Grid Security Testing and Standardization. In Proceedings of the 2020 IEEE International Workshop on Metrology for Industry 4.0 & IoT, Roma, Italy, 3–5 June 2020; pp. 435–440. [CrossRef]
33. Castellani, A.; Schmitt, S.; Squartini, S. Real-world Anomaly Detection by Using Digital Twin Systems and Weakly Supervised Learning. *IEEE Trans. Ind. Inform.* **2021**, *17*, 4733–4742. [CrossRef]
34. Murillo, A.; Taormina, R.; Tippenhauer, N.; Galelli, S. Co-simulating Physical Processes and Network Data for High-fidelity Cyber-security Experiments. In Proceedings of the Sixth Annual Industrial Control System Security (ICSS) Workshop, 2020, ICSS 2020, Austin, TX, USA, 8 December 2020; pp. 13–20. [CrossRef]
35. Saad, A.; Faddel, S.; Mohammed, O. Iot-based Digital Twin for Energy Cyber-physical Systems: Design and Implementation. *Energies* **2020**, *13*, 4762. [CrossRef]
36. Suhail, S.; Jurdak, R.; Matulevicius, R.; Seon Hong, C. Securing Cyber-physical Systems through Blockchain-based Digital Twins and Threat Intelligence. *arXiv* **2021**, arXiv:2105.08886.
37. Chukkapalli, S.S.L.; Pillai, N.; Mittal, S.; Joshi, A. Cyber-physical System Security Surveillance Using Knowledge Graph Based Digital Twins—A Smart Farming Usecase. In Proceedings of the 2021 IEEE International Conference on Intelligence and Security Informatics (ISI), Antonio, TX, USA, 2–3 November 2021; pp. 1–6. [CrossRef]
38. Danilczyk, W.; Sun, Y.L.; He, H. Smart Grid Anomaly Detection Using a Deep Learning Digital Twin. In Proceedings of the 2020 52nd North American Power Symposium (NAPS), Tempe, AZ, USA, 11–13 April 2021; pp. 1–6. [CrossRef]
39. Patel, A.; Schenk, T.; Knorn, S.; Patzlaff, H.; Obradovic, D.; Halblaub, A.B. Real-time, Simulation-based Identification of Cyber-security Attacks of Industrial Plants. In Proceedings of the 2021 IEEE International Conference on Cyber Security and Resilience (CSR), Virtual, 26–28 July 2021; pp. 267–272. [CrossRef]
40. Garcia, H.E.; Aumeier, S.E.; Al-Rashdan, A.Y.; Rolston, B.L. Secure Embedded Intelligence in Nuclear Systems: Framework and Methods. *Ann. Nucl. Energy* **2020**, *140*, 107261. [CrossRef]
41. Tärneberg, W.; Skarin, P.; Gehrmann, C.; Kihl, M. Prototyping Intrusion Detection in an Industrial Cloud-native Digital Twin. In Proceedings of the International Conference on Industrial Technology, Valencia, Spain, 10–12 March 2021.

42. Dietz, M.; Englbrecht, L.; Pernul, G. Enhancing Industrial Control System Forensics Using Replication-based Digital Twins. In *Advances in Digital Forensics XVII*; Peterson, G., Shenoi, S., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2021; Volume 612, pp. 21–38. [CrossRef]
43. Dietz, M.; Schlette, D.; Pernul, G. Harnessing Digital Twin Security Simulations for systematic Cyber Threat Intelligence. In Proceedings of the 2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC), Los Alamitos, CA, USA , 27 June–1 July 2022; pp. 789–797. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.