applied sciences

MDPI

Article

Building Cyber-Resilient Smart Grids with Digital Twins and Data Spaces

Luigi Coppolino [1,*,†], Roberto Nardone [1,†], Alfredo Petruolo [1,†] and Luigi Romano [2,*,†]

[1] Department of Engineering, University of Naples "Parthenope", 80143 Naples, Italy; roberto.nardone@uniparthenope.it (R.N.); alfredo.petruolo001@studenti.uniparthenope.it (A.P.)
[2] Department of Economic, Legal, IT and Motor Sciences, University of Naples "Parthenope", 80143 Naples, Italy
* Correspondence: luigi.coppolino@uniparthenope.it (L.C.); luigi.romano@uniparthenope.it (L.R.)
† These authors contributed equally to this work.

Abstract: The rapid expansion of digital twin technology has revolutionized management and testing across various sectors, particularly in safeguarding critical infrastructure like smart grids. Aligned with the NIS2 Directive, Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022, digital twins play a critical role in bolstering cybersecurity measures by enabling advanced simulation and real-time monitoring, key aspects emphasized in the directive for enhancing the security of networks and information systems. Concurrently, Europe's shift towards communal data spaces amplifies the need for robust cyber defences. This paper aims to bolster cybersecurity defences in critical infrastructure, with a particular focus on the energy sector and smart grids. It proposes an innovative architecture for cybersecurity monitoring, converting a Common Information Model-compliant system into a digital twin via the FIWARE platform and incorporating an open-source Security Information and Event Management solution. Validated by a real-world case study, our approach demonstrates significant advancements in protecting smart grids against cyber threats.

Keywords: EPES; FIWARE; SIEM; cybersecurity; monitoring systems

Citation: Coppolino, L.; Nardone, R.; Petruolo, A.; Romano, L. Building Cyber-Resilient Smart Grids with Digital Twins and Data Spaces. *Appl. Sci.* **2023**, *13*, 13060. https://doi.org/10.3390/app132413060

Academic Editor: Gianluca Lax

Received: 7 November 2023
Revised: 3 December 2023
Accepted: 3 December 2023
Published: 7 December 2023

1. Introduction

In recent years, the imperative to protect critical infrastructure has become a paramount concern, driven by both industry and academia. As the digital transformation paves the way for more sustainable and efficient services, it simultaneously brings to the fore the essentiality of stringent cybersecurity measures. This is particularly true in the energy sector, where the integration of cyber-physical systems has introduced a spectrum of new security vulnerabilities and safety concerns. The European Commission's introduction of the NIS2 directive [1] marks a significant step towards reinforcing the resilience and defence of these vital systems. In this landscape, the deployment of digital twins has emerged as a crucial strategy for the proactive assessment of cybersecurity risks [2–4]. These digital replicas of the physical world provide a platform for vigilant monitoring, enabling the early detection and identification of threats as well as the introduction of remediation strategies into critical infrastructure.

Digital twins also enable organizations to examine different scenarios, assess the robustness and effectiveness of security measures, and identify areas for improvement. Such a proactive approach is indispensable for shielding critical infrastructures from the dynamic spectrum of the ever-changing landscape of cyber threats. Furthermore, the European Commission has also highlighted the strategic role of common data spaces in achieving data sovereignty, promoting secure data sharing across different sectors, and encouraging stakeholder collaboration [5,6]. In the energy sector, the establishment of a shared energy data space could lead to improved citizen services and better management systems. Among the different initiatives, FIWARE has emerged as a prominent open-source

platform that facilitates the creation of smart applications, ensuring interoperability and fostering an environment of open collaboration [7]. Nonetheless, the Common Information Model (CIM) [8] continues to be the gold standard for data exchange within the power system industry.

Bridging the established CIM standard with emerging common data spaces is essential to amplify insights, bolster decision making, and elevate the operational efficiency in the energy sector. This confluence is poised to unlock unprecedented opportunities, catalyze stakeholder cooperation, and accelerate advancements in energy management and service delivery.

This paper introduces a framework that leverages the capabilities of FIWARE to tackle the challenge of cybersecurity monitoring in smart power and distribution grids. FIWARE is pivotal in our research, offering a versatile platform that aligns seamlessly with the Common Information Model for smart grid cybersecurity. Its robust support for data spaces and digital twin integration underpins our approach, enhancing interoperability and real-time data analysis. This alignment with FIWARE significantly amplifies the effectiveness of our cybersecurity solutions in smart grid environments. This framework is designed to facilitate the development of data-driven digital twin applications, built upon open-source platforms while maintaining adherence to established industry standards. The ultimate objective is to enhance the cybersecurity monitoring of smart grids through a digital twin that accurately tracks the overall status of the CIM-based architecture of the real system. Enhancing advanced real-time data processing and scalable search capabilities is crucial for efficiently managing and analyzing large-scale, complex datasets in cyber threat detection and response. In our framework, OpenSearch has been integrated to support this activity. Leveraging OpenSearch is necessary in the framework to support the real-time data analysis and scalable search capabilities, essential for smart grid cybersecurity. Its ability to handle complex, voluminous datasets aligns perfectly with the demands of monitoring and protecting smart grid systems. OpenSearch thus plays a key role in our architecture, offering enhanced efficiency and precision in detecting and responding to cyber threats.

Our proposed architecture, while specifically utilizing technologies like FIWARE and OpenSearch, is general enough and can be adapted to different technological contexts. The selected technologies were chosen for their current, cutting-edge capabilities, demonstrating the practical implementation of our architecture. However, this design maintains enough flexibility to be applicable and effective with a range of similar future technologies, ensuring its long-term relevance and adaptability. It extends the research initiated in [9], advancing towards a more sophisticated and resilient methodology. The main improvement is to augment the existing functionalities to be in harmony with the principles of data spaces, with a particular emphasis on incident-reporting protocols that conform to IDMEF v2 standards. This improvement also enables the immediate reaction to the discovered threats, orchestrating the needed actions and communicating with the involved external actors. The proposed enhancement strictly adheres to the European vision of constructing an energy data space and could serve as a foundation for collaborative future research.

The effectiveness of such a framework is demonstrated through its application to a real-world power distribution grid in Kropa, Slovenia. In the proposed case study, a digital twin was constructed starting from the CIM-based representation of the system, while the real-time monitoring and reaction strategies are facilitated by the SIEM solution based on open-source technologies like OpenSearch. The Kropa grid's case study provided a practical context to evaluate the digital twin's monitoring capabilities, allowing for the analysis of data, detection of irregularities, and obtention of insights into the system's performance and security.

The rest of the paper is organized as follows. Section 2 introduces the concept of a digital twin and data spaces, which are the basics of the current research. Section 3 gives all the details about the proposed framework. Section 4 validates the proposal against the considered real-world case study. Section 5 analyzes the current research on the topic

and compares the proposal with it. Section 6 draws some conclusions and addresses the future work.

2. Digital Twins and Data Spaces

This section introduces the concepts of digital twins and data spaces. The integration of these two concepts is at the basis of our proposal, where digital twins utilize the data within data spaces to create accurate and dynamic models, enabling predictive analytics and informed decision-making.

2.1. Digital Twin

A digital twin (DT) is a virtual replica of a physical system, created using sensors, data analytics, and machine learning algorithms. It captures and analyzes real-time data from the physical system and has the potential to transform the monitoring, analysis, and optimization of physical objects and systems in Industry 4.0 [10–12]. Digital twins allow engineers to gain insights from real-time data generated by IoT devices, enabling simulation, analysis, and optimization for increased efficiency, productivity, and cost savings.

There are three types of digital twins:

- Digital Twin Prototype. This type underpins the conceptualization and development phase of a physical asset, serving as a blueprint for its design.
- Digital Twin Instance. This variant maintains a continuous digital synchronicity with its physical counterpart, evolving over the asset's lifecycle.
- Digital Twin Aggregate. This type amalgamates data from multiple DT instances to conduct comprehensive simulations and forecasts pertinent to the overarching physical entity.

The construction of a digital twin involves several key phases:

1. Data Acquisition (DA): This phase involves the collection of real-time data from the physical system (PS) using a network of sensors (S). The function for data acquisition can be defined as:
$$DA : S \rightarrow D_{rt}(PS) \tag{1}$$
where $D_{rt}(PS)$ denotes the real-time data stream from the physical system.
2. Modeling (M): In this phase, real-time data D_{rt} are utilized to construct a virtual model (VM) that represents the physical system. The modelling function can be expressed as:
$$M : D_{rt} \rightarrow VM \tag{2}$$
which maps the real-time data to a corresponding virtual model.
3. Simulation (Sim): Using the virtual model, this phase performs simulations to emulate the behaviour (B) of the physical system under various conditions (C). The simulation function is:
$$Sim : VM \times C \rightarrow B \tag{3}$$
where B represents the behaviours or outcomes under simulated conditions.
4. Deployment (Dep): The final phase involves integrating the digital twin (DT) with the physical system to enable ongoing optimization (Opt). The deployment function can be formalized as:
$$Dep : (DT, PS) \rightarrow Opt \tag{4}$$
which indicates that the DT and the physical system are integrated to achieve an optimized state.

Each function represents a mapping from inputs to outputs, reflecting the transformation of data and models throughout the DT building process.

2.2. Data Spaces

The European data strategy is geared towards the creation of a single, unified data market to boost Europe's competitive edge while preserving its data sovereignty. This strategy encourages the establishment of shared European data spaces, which are designed to enhance data accessibility for economic and societal benefits while ensuring that data producers, whether businesses or individuals, retain control and rights over their data.

The management of the European data space adheres to core European values and complies with pertinent legislation, such as the General Data Protection Regulation (GDPR) and the recently updated Network and Information Systems Directive (NIS2). These regulations are pivotal in guaranteeing a secure and seamless data flow within a safeguarded ecosystem. Moreover, initiatives like the International Data Spaces Association (IDSA), GAIA-X, and FIWARE Foundation are pivotal in driving the adoption of data-centric technologies and establishing data sovereignty.

FIWARE has also emerged as a key platform that propels the development and uptake of innovative, data-centric technologies. It is an open-source platform that comes equipped with an extensive suite of standards, APIs, and tools tailored for crafting and managing intelligent applications and services. FIWARE's main objectives are to ensure the interoperability, reusability, and scalability of applications across a variety of sectors.

With FIWARE, developers are empowered to devise and implement applications that leverage real-time data from diverse sources, such as sensors, IoT devices, and pre-existing systems. It provides a sturdy infrastructure for data handling, processing, and analysis, and also supports smooth integration with cloud services and other external platforms. Utilizing FIWARE enables developers to create state-of-the-art applications that exploit the full potential of real-time data, fostering innovation and operational efficiency across various industries.

To encapsulate the concept of the data space ($\mathcal{DS}$), in a FIWARE-compliant view, we introduce the following notation:

$$\mathcal{DS} = \langle D, P, S, St, G \rangle \tag{5}$$

where:

- $D = \{d_1, d_2, \ldots, d_n\}$ represents a set of datasets.
- $P = \{p_1, p_2, \ldots, p_m\}$ denotes a set of policies governing the data space.
- $S = \{s_1, s_2, \ldots, s_k\}$ is a set of services that facilitate data processing and analysis.
- $St = \{st_1, st_2, \ldots, st_j\}$ indicates a set of stakeholders involved in the data space.
- $G : \mathcal{S} \rightarrow \{compliant, non\text{-}compliant\}$ is a governance function for the data space that is responsible for evaluating the compliance of services with the established regulatory and policy framework of the data space.

This function examines each service to ensure it adheres to regulations such as the General Data Protection Regulation (GDPR) and the Data Act, maintaining secure and sovereign data operations.

The following Algorithm 1 evaluates whether a data service adheres to the Data Act's regulatory framework by conducting a streamlined compliance check.

Algorithm 1 Data space service compliance

```
procedure CheckPrincipleCompliance(Service, Principle)
    isCompliant ← AssessCompliance(Service, Principle)
    if not isCompliant then
        return false                ▷ Service does not comply with the principle
    else
        return true                 ▷ Service is compliant with the principle
    end if
end procedure
```

Algorithm 1 *Cont.*

```
function AssessCompliance(Service, Principle)
    if Principle == Data Sharing Principle then
        return Service.meetsDataSharingRequirements()
    else
        if Principle == Data Transparency Principle then
            return Service.ensuresDataTransparency()
        else
            if Principle == Data Portability Principle then
                return Service.supportsDataPortability()
            else
                return false                          ▷ Principle not recognized
            end if
        end if
    end if
end function
```

3. Proposed Framework

This section describes the architecture of the proposed framework, together with a comprehensive overview of the Common Information Model and data spaces in smart grids.

3.1. Common Information Model—IEC 61970

The Common Information Model (CIM) [8] is an essential standard that serves as the backbone for the integration and management of information systems in the power system sector. It offers a uniform structure for data representation and exchange, enhancing compatibility and enabling smooth communication across the various segments of power systems. The drive for a standardized information model emerged from the increasing intricacy and heterogeneity of electrical networks. As the energy sector evolved, a plethora of standalone software solutions and devices emerged, leading to isolated data repositories and suboptimal data flow. This fragmentation became a significant hurdle when it came to incorporating new technologies such as renewables, smart grids, and demand response initiatives. Originating from the efforts of the Electric Power Research Institute (EPRI) in the 1990s, the CIM was designed to forge a common language for electrical system representation. It later gained international stature, being refined by the International Electrotechnical Commission (IEC) and the International Organization for Standardization (ISO). The CIM delineates an extensive, flexible schema of data models that articulate the elements and interconnections characteristic of the power system landscape. These models span a broad spectrum, from power production and delivery to market transactions and customer engagement. By unifying the depiction of these elements, the CIM paves the way for different systems and applications to work in concert. A principal benefit of the CIM is its provision of an integrated perspective of the energy network, allowing for a more cohesive and efficient management of a power system.

The CIM harnesses the Unified Modeling Language (UML) [13] to graphically articulate the structure and behaviour of its system components [14]. UML's universally recognized graphical tools offer a suite of diagrams and symbols that deliver a coherent visual representation of the CIM's constructs. This includes a variety of diagrams such as class, object, and sequence diagrams, each capturing distinct facets of the CIM. Through UML, the CIM's entities, with their respective attributes and interrelations, are depicted, providing an intelligible and succinct map of the model's architecture.

Adapting to the dynamic requirements of the power system domain, the CIM has expanded to incorporate modern technological practices, including web services and XML-based data exchange protocols, to maintain relevance and ensure seamless integration with current IT frameworks. A significant stride in this evolution is the FIWARE community's integration of CIM entities into its data models. This harmonization between the CIM

and FIWARE data models is instrumental in enabling efficient data interchange between systems compliant with CIM standards and those developed on the FIWARE architecture, fostering an ecosystem where data circulate freely and efficiently.

3.2. Data Space for Smart Grids

In the face of escalating energy demands driven by economic growth, industrial expansion, and population increases, the quest for sustainable energy solutions has become increasingly critical. Industries such as manufacturing, mining, and transportation, in addition to the development of urban infrastructures, are experiencing a steep climb in energy consumption [15]. The reliance on fossil fuels to satisfy these needs, however, is a cause for environmental concern, as their combustion emits greenhouse gases, exacerbating global warming and climate change [16]. Addressing these challenges, the modernization of power grids emerges as a pivotal strategy, transitioning from traditional systems hampered by an ageing infrastructure and a limited renewable integration capacity to smart grids that promise enhanced efficiency, resilience, and renewable compatibility [17].

The European Commission's initiative to create common data spaces in the energy sector is a testament to this transition, with smart grids at the forefront, enabling a two-way flow of energy and data that enriches stakeholder insights. Among the supported initiatives, FIWARE is notable for its role in shaping these data spaces through standardized data models and generic enablers, which are instrumental in developing collaborative and smart solutions.

Central to FIWARE's ecosystem is the Context Broker, a key component that facilitates the creation of digital twins, essential for smart grid monitoring. It allows for the collection and analysis of data, improving grid management through its subscription methods. FIWARE's architecture also addresses interoperability issues by incorporating IoT agents that act as conduits for data exchange, converting various messaging protocols into a unified FIWARE-compliant format. This harmonization ensures that different devices and systems within the smart grid can communicate effectively. The FIWARE architecture, depicted in Figure 1, encapsulates the elements and interactions within this ecosystem, as discussed in this section.

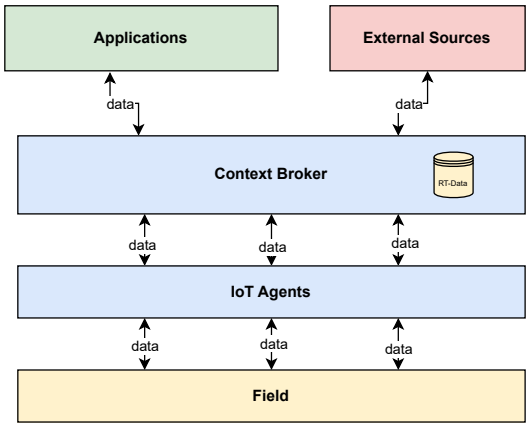


Figure 1. FIWARE core components.

3.3. Framework Architecture

The architectural design we propose for our framework is tailored to enhance the surveillance of essential services, particularly focusing on advanced smart grid systems. This design incorporates various elements as illustrated in Figure 2. Our process starts with the transformation of smart grid data, organized using the CIM, into the NGSI-LD format, a standard designed for interoperability and compliance with the data space view. These data are then relayed to the Context Broker, a digital mirror of the physical grid that oversees the grid's entities, their relationships, and their data attributes. The Context Broker is equipped with subscription features that notify the system of any changes in the

grid's condition. These notifications are sent to the Kafka message broker, which archives historical data and facilitates interaction with external applications. Figure 2 also depicts the main technologies that have been used in each module, which will be described in detail in the following paragraph.

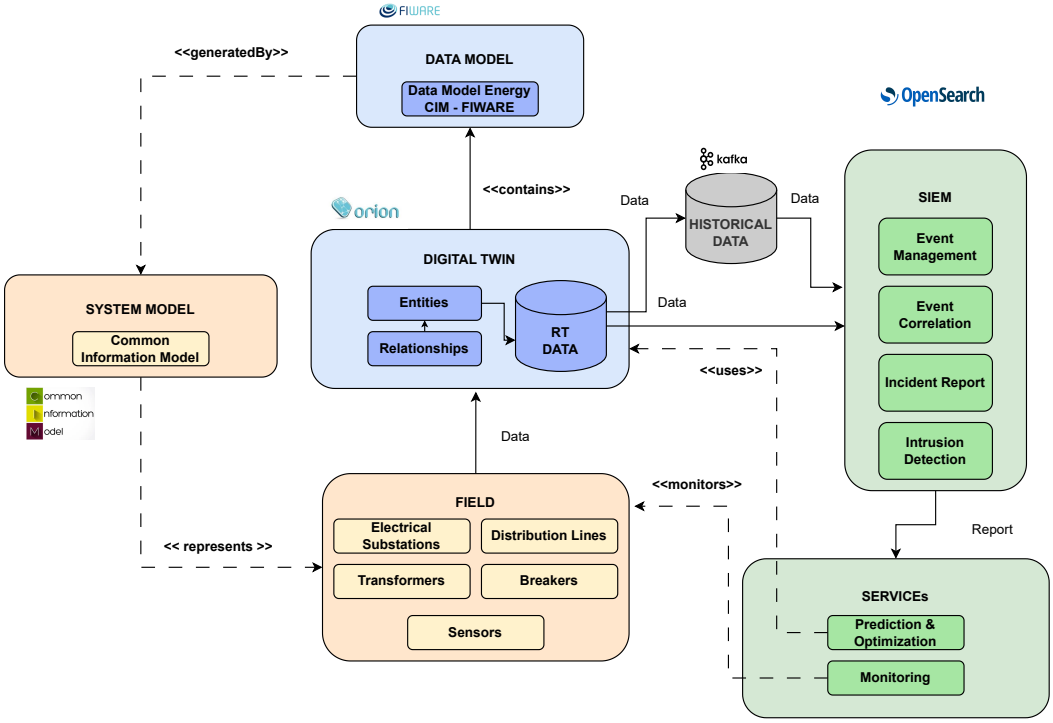


Figure 2. Framework architecture.

From the CIM to the data space. The initial step in the concrete implementation of the proposed architecture involves the critical process of data conversion. Data originating from the field, already modelled following the CIM standard, are transformed for the FIWARE Context Broker. This step is pivotal, not only for aligning with the stringent regulations governing data spaces but also for guaranteeing the uniformity of data across the system. It paves the way for a standardized data model that facilitates seamless integration and interoperability among the diverse array of applications and services deployed within this ecosystem. FIWARE's suite of data models caters to various sectors, offering specialized solutions tailored to each domain. Pertinently, within the energy domain, we leverage the CIM-compliant data model provided by FIWARE, which we have identified as the most suitable for our case study. The CIM is typically implemented in XML or RDF syntaxes, which are both highly structured file formats. These formats are designed to be comprehensive and extensible, capable of representing the complex relationships and attributes of electrical systems. However, the level of detail and the strict schema adherence required can make CIMs somewhat rigid and cumbersome for rapid integration and real-time data exchange.

On the other hand, the NGSI-LD (Next-Generation Service Interface–Linked Data) model is a modern and context-aware API specification for context information management. It builds upon the legacy of NGSI v2 by incorporating linked data, which enables the data to be fully interoperable and machine-understandable. NGSI-LD uses a JSON-LD format, which is inherently more flexible and web-friendly than the XML/RDF formats typically used with the CIM. This flexibility allows NGSI-LD to support more dynamic and real-time applications, making it well-suited to smart applications that require rapid processing and integration of data from various sources.

The main differences between CIM and NGSI-LD are reported in Figure 3, and can be summarized as follows:

- Syntax and Format: CIM's XML/RDF formats are document-centric and can be more verbose, whereas NGSI-LD's JSON-LD format is data-centric, lighter in weight, and more conducive to web transmission.
- Flexibility: The CIM is schema-driven and can be less flexible in accommodating changes, whereas NGSI-LD is designed to be more adaptable to the evolving needs of smart applications.
- Interoperability: While CIM ensures interoperability within the energy sector, NGSI-LD's use of linked data principles extends interoperability across different domains and applications.
- Real-time Processing: NGSI-LD is optimized for real-time context awareness and processing, which is essential for dynamic environments like smart cities, whereas the CIM's traditional use cases involve more static data exchanges.

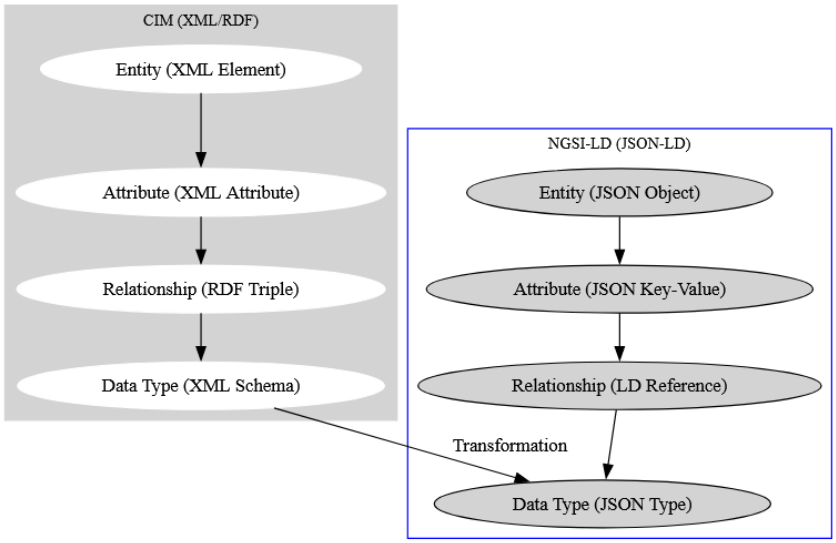


Figure 3. Comparative visualization of CIM (XML/RDF) and NGSI-LD (JSON-LD) syntax structures.

In the context of smart grids, converting CIM-formatted data into NGSI-LD format is a transformative step that aligns the static, heavily modeled data with the dynamic, real-time operational needs of modern energy systems. This conversion facilitates the transition from a traditional grid management approach to a more agile and intelligent smart grid paradigm, where data are not only exchanged but also understood and utilized in a context-aware manner.

Orion DT and Data Management. The construction of the digital twin is facilitated through the FIWARE Orion Context Broker, which serves as the central hub for managing real-time contextual information about the infrastructure. Within this framework, entities are meticulously defined by specified data models and are enriched with detailed interrelationships. The Orion Context Broker operates via a RESTful API, allowing for operations to be executed through standard HTTPS methods. Table 1 provides a comprehensive summary of the available operations along with their corresponding expected payloads.

Table 1. Summary of Orion Context Broker operations.

HTTP Method	Operation	Description
GET	Retrieve Entity	Retrieves the current state of an entity based on its ID.
POST	Create Entity	Creates a new entity with the provided attributes and values.
PUT	Update Entity	Updates an existing entity with new or modified attributes.
PATCH	Partial Update	Partially updates attributes of an existing entity.
DELETE	Remove Entity	Deletes an entity from the Context Broker.

Regarding data stewardship, FIWARE incorporates MongoDB, which lays the groundwork for preserving the system's real-time data integrity. The segregation of real-time and historical data is a strategic approach that enhances the performance and efficiency of data management systems, especially in the context of a digital twin. Real-time data are dynamic and constantly changing, and they require immediate processing to accurately reflect the current state of the physical infrastructure. These data are typically transient and are stored in high-performance databases like MongoDB, which the FIWARE Orion Context Broker uses. Such databases are optimized for speed and low-latency access, ensuring that the digital twin is a real-time representation of the physical counterpart.

Historical data, on the other hand, consist of time-series records of past states and events. These data are invaluable for trend analysis, predictive maintenance, and long-term strategic planning. Storing these data requires a different approach, as the emphasis is on data retention, query flexibility, and cost-effective scalability rather than immediate write and read access. The need to separate real-time data from historical data stems from performance considerations. Real-time data systems prioritize a low latency to ensure that the digital twin is synchronized with the physical infrastructure. In contrast, historical data systems are optimized for cost-effective storage and complex data retrieval operations, which are typically more resource-intensive and can introduce latency if mixed with real-time processing.

By maintaining this separation, organizations can ensure that the performance of the digital twin remains unaffected by the intensive computational demands of historical data analysis.

External Application Interaction. The architecture we propose streamlines the incorporation of external applications by leveraging a uniform data model and the digital twin's comprehensive representation of the smart grid. In our security-centric approach to safeguarding the smart grid infrastructure, we have adopted OpenSearch, an open-source SIEM (Security Information and Event Management) platform. OpenSearch capitalizes on the standardized data format provided by the digital twin, which is systematically archived within the Kafka broker. This harmonization of data is crucial as it ensures that OpenSearch can efficiently process and analyze information without the need for extensive data normalization procedures. By interfacing directly with the Kafka broker, OpenSearch can access a continuous stream of real-time and historical data, enabling a suite of cybersecurity services that is essential for robust smart grid protection.

As highlighted in Figure 1, the services facilitated by OpenSearch include:

- Event Management: OpenSearch aggregates and manages events across the smart grid, providing a centralized view of all activities. This enables operators to monitor the grid's operational status and detect anomalies in real time.
- Event Correlation: By correlating disparate events, OpenSearch can identify patterns that may indicate complex cyber threats. This correlation is vital for recognizing multi-stage attacks that single events might not reveal.
- Incident Reporting: OpenSearch automates the generation of incident reports, which are critical for compliance with regulatory standards and for initiating appropriate response procedures.
- Intrusion Detection: Utilizing advanced analytics and pattern recognition, OpenSearch serves as an intrusion detection system, spotting potential security breaches and alerting operators to take pre-emptive measures.

Through the integration of OpenSearch with the digital twin, our architecture not only enhances the security posture of smart grids but also paves the way for advanced analytical capabilities. This integration allows for the proactive management of cybersecurity threats, ensuring the resilience and reliability of energy systems in the face of evolving cyber risks. Table 2 reports the capabilities of OpenSearch in these activities.

Table 2. Capabilities of OpenSearch in smart grid cybersecurity.

Service	Function	Benefit
Event Management	Aggregates and manages events, providing a centralized activity view	Enables real-time monitoring and anomaly detection
Event Correlation	Correlates events to identify complex cyber threat patterns	Aids in recognizing multi-stage attacks
Incident Reporting	Automates incident report generation	Ensures compliance and initiates response procedures
Intrusion Detection	Utilizes analytics for threat detection	Alerts operators to potential security breaches

4. Case Study

To ascertain the efficacy of our architectural design, we conducted a case study on an operational smart grid in Kropa, a Slovenian village. Table 3 provides a comprehensive depiction of the infrastructure, shedding light on the individual components that make up the grid. Our empirical investigation pursued two primary objectives: Firstly, to conduct a performance analysis of the architectural elements employed, with a specific focus on standard CRUD operations to assess the digital twin's efficiency. We measured the latency and throughput of these operations under various load conditions, simulating the concurrent activities of multiple grid elements. The results indicated that the digital twin responded within acceptable time frames, even under peak load, thus validating its scalability and robustness.

Secondly, we aimed to demonstrate the SIEM system's proficiency in detecting potential faults within the grid and generating automated reports for the Security Operations Center, thereby enabling alerts in the event of a system failure. To this end, we introduced a series of controlled fault conditions into the system to observe the SIEM's detection and reporting capabilities. The SIEM system successfully identified and categorized each fault, triggering automated alerts that were then verified for accuracy and timeliness by the SOC team. These experiments confirmed the system's capability to serve as an early warning mechanism, potentially preventing minor issues from escalating into major outages.

Finally, the deployment of both the Orion Context Broker and the SIEM was carried out on a laptop configured with an Intel® Core™ i7-1270P processor and 32 GB of LPDDR5 RAM at 4800 MHz.

Table 3. Overview of electrical network components in Kropa's smart grid.

Component	Instances	Description
Substation	30	Key facilities for voltage transformation using transformers.
PowerTransformer	8	Devices that transfer electrical energy between voltage levels.
EnergyConsumer	41	Loads or consumers of electricity, such as residential or commercial customers.
ACLineSegment	80	Transmission lines that carry high-voltage electricity over distances.
Breaker	26	Circuit breakers that interrupt current flow after fault detection.
Disconnector	40	Switches for isolating parts of the network for maintenance or safety.
BusbarSection	36	Used within substations to distribute power and manage connections.

4.1. Threat Modeling and Attack Scenario

In previous work, we have outlined the critical elements that form the backbone of a smart grid system, each essential for the uninterrupted flow and control of electricity. Yet, the very interconnectivity that enables such efficiency also opens the door to a host of vulnerabilities that could be targeted by adversaries. A forward-looking and holistic approach to threat modeling is crucial for this infrastructure, taking into account both the tangible elements and the digital interfaces that command them.

Key infrastructure points such as substations, power transformers, and various consumer connections are at risk of cyber-physical attacks that may grant unauthorized parties control over these systems. Adversaries could potentially exploit gaps in the network's

defences or use targeted phishing schemes to gain entry into the operational technology systems. Once they breach the perimeter, they could establish backdoors or alter the configuration of essential components, which might trigger a chain reaction of instability across the grid. The consequences of such breaches are dire, ranging from blackouts to physical damage, economic losses due to interrupted services, and significant safety risks for both the public and workers. In Figure 4, we present a NIST attack tree that outlines the sequential phases of a sophisticated cyber attack targeting the smart grid infrastructure. This visual representation captures the methodical approach taken by attackers, from initial intelligence gathering to the final destabilization of the grid.

Attack Scenario. The attack scenario unfolds as a calculated and stealthy operation aimed at undermining the integrity of the smart grid infrastructure. It begins with the meticulous gathering of open-source intelligence, where attackers compile detailed information about the grid's layout and the personnel who operate it. This intelligence forms the foundation for a spear-phishing campaign, tailored to deceive and exploit employees with access to critical control systems.

Once the phishing attack succeeds, the intruders establish a backdoor, granting them clandestine access to the network. This unauthorized entry point is the precursor to a series of insidious actions, culminating in the manipulation of the substation's control settings. The attackers' ultimate goal is to induce grid instability, triggering a domino effect of power disruptions and system failures.

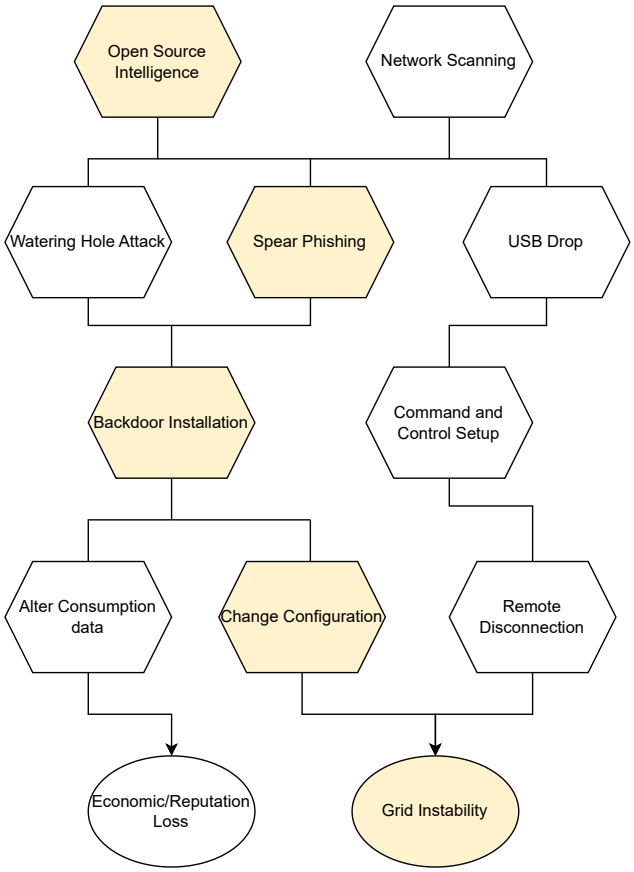


Figure 4. Attack tree on Kropa's infrastructure.

4.2. SIEM in the Case Study

In light of the presented attack scenario, our architectural framework is designed with a strong emphasis on advanced monitoring capabilities to safeguard the infrastructure against cyber threats. Central to this design is the deployment of OpenSearch, which serves as the backbone for real-time data analysis and operational oversight of the smart grid.

The configuration change phase, a critical juncture in the attack scenario, is addressed with OpenSearch's ability to closely monitor and visualize operational parameters. This is achieved through a dynamic dashboard, built using Graphana, that reflects real-time data, enabling the detection of unauthorized configuration changes. The system is configured to recognize deviations from standard operational parameters, ensuring that any abnormal activity is immediately flagged. Within the scope of our research, we have developed a dashboard, illustrated in Figure 5, that displays real-time telemetry data from electrical substations. The dashboard is engineered to parse and render time-series data, capturing a range of electrical parameters that are critical to the operational integrity of power distribution systems. The data include precise timestamps (coordinated universal time—UTC), substation identifiers, electrical loads in megawatts (MW), voltage levels in kilovolts (kV), the system frequency in hertz (Hz), the ambient temperature in degrees Celsius (°C), and the operational status of each substation.

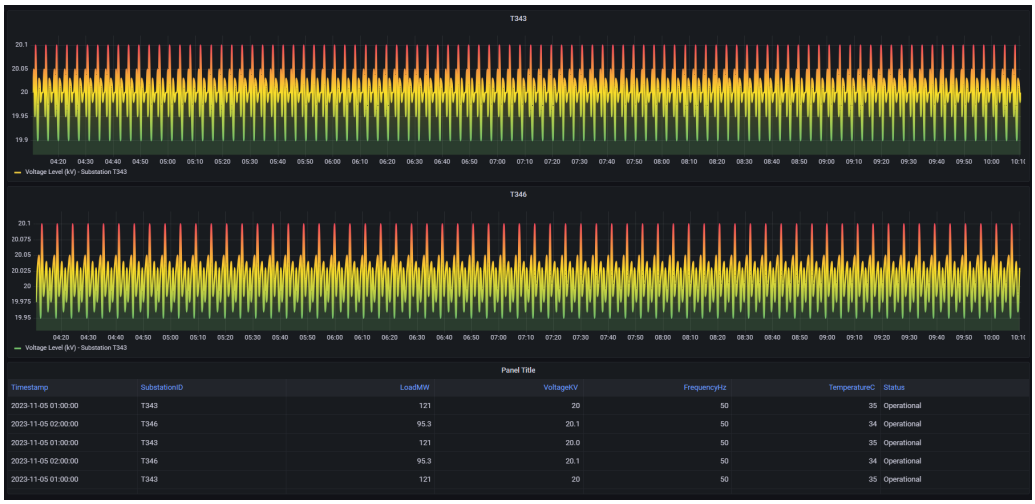


Figure 5. Real-time dashboard display of electrical substation parameters.

Furthermore, the system architecture supports the implementation of IDMEF (Intrusion Detection Message Exchange Format) version 2 [18], allowing for standardized incident reporting and facilitating interoperability between our monitoring systems and external cybersecurity frameworks. The report generated below exemplifies this functionality, detailing an anomalous voltage oscillation at substation T343. Such reports are vital for the prompt diagnosis and rectification of issues, thereby ensuring the reliability and safety of the power distribution network. Moreover, our reporting structure is fully compliant with the IDMEF version 2 standards, which facilitates interoperability and standardized communication within cybersecurity management processes. An example of such an event is given in the following Listing 1.

The strategic incorporation of OpenSearch into our framework significantly enhances the operational visibility of the smart grid. It also strengthens the reactive capabilities of our security measures. By providing a comprehensive view of the grid's health and enabling quick detection and communication of potential threats, our approach solidifies the grid's defence against sophisticated cyber attacks, thereby ensuring its continuous and secure operation.

```
{
  "Version": "2.0.3",
  "ID": "a1b2c3d4-5678-90ab-cdef-1234567890ab",
  "Entity": "Substation T343",
  "Category": [
    "System.Anomaly.VoltageOscillation"
  ],
  "Cause": "Technical Fault",
  "Description": "Anomalous voltage oscillation",
  "Status": "Incident",
  "Severity": "High",
  "CreateTime": "2023-11-05T12:34:56Z",
  "DetectTime": "2023-11-05T12:34:55Z",
  "Confidence": 0.95,
  "Analyzer": {
    "IP": "192.168.1.10",
    "Name": "Voltage Monitoring System",
    "Hostname": "volt-mon.example.com",
    "Type": "System",
    "Category": [
      "SCADA"
    ],
    "Data": [
      "Voltage Readings"
    ],
    "Method": [
      "Automated Monitoring"
    ],
  "Target": [
    {
      "Location": "Power Grid"
    }
  ],
  "Vector": [
    {
      "Category": "System",
      "Name": "Substation T343",
      "Size": "Large",
      "Observable": [
        "Voltage Oscillation"
      ]
    }
  ],
  "Observables": [
    {
      "Name": "Voltage Oscillation",
      "Reference": "IDMEFv2",
      "Content": "{\"type\":\"Voltage\",\"Value\":\"20.150kV\"}"
    }
  ]
}
```

Listing 1. Example of a report in IDMEF format.

4.3. Performance Analysis

In the smart grid context, the responsiveness and scalability of the digital twin are critical, especially for applications, such as smart grids, where real-time data processing is paramount. To this end, we have conducted an in-depth performance analysis of the FIWARE Orion Context Broker, which is crucial in managing the real-time state of digital twins in a smart grid environment. This experiment was designed to simulate a scenario where a multitude of IoT devices, akin to sensors and actuators within a smart grid, concurrently send requests to the Context Broker, thereby mimicking the real-world operational load and data throughput.

The Orion Context Broker was deployed on a local server, and the experiment was configured to measure the latency of CRUD operations. The CRUD operations are defined as follows:

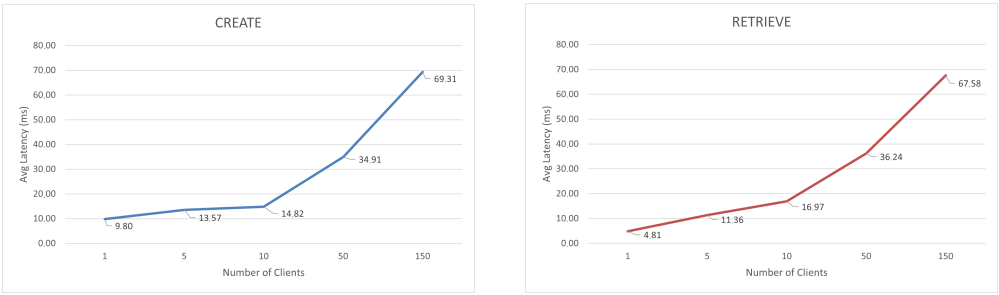
- Create: Register a new entity with a unique ID and predefined attributes.
- Read: Retrieve the entity's details using its unique ID.
- Update: Modify the entity's attributes.
- Delete: Remove the entity from the Context Broker.

These operations are fundamental to the interaction between IoT devices and the Context Broker, reflecting the typical transactional requirements within a smart grid infrastructure. The performance of these operations directly impacts the efficiency and reliability of the digital twin representation and, by extension, the operational integrity of the smart grid itself.

To emulate the concurrent operations from multiple IoT devices, we utilized a Python script that employs the `'concurrent.features'` module, allowing us to simulate up to 150 clients interacting with the Context Broker simultaneously. Each simulated device performed a sequence of CRUD operations, with a total of 50 requests per operation type to ensure a comprehensive assessment. The script recorded the time taken for each operation, and these metrics were aggregated to calculate the average latency for each CRUD operation across varying levels of client concurrency.

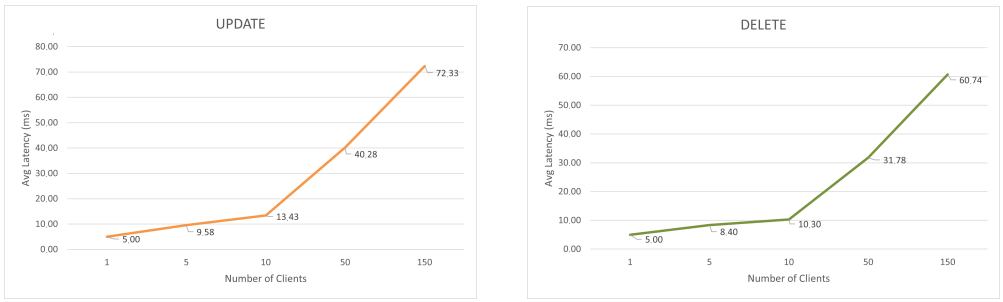
The significance of this experiment lies in its realistic simulation of IoT device behaviour, providing a robust model for understanding how the Orion Context Broker performs under different loads. This is particularly relevant for smart grid applications where the timely processing of data from numerous sources is essential for maintaining grid stability and responding to dynamic conditions.

The results, shown in Figure 6, are useful for identifying performance bottlenecks and optimizing the Context Broker's configuration for real-world deployment.



(**a**) (**b**)

Figure 6. *Cont*.



(**c**) (**d**)

Figure 6. Performance analysis of CRUD operations at varying levels of client concurrency. (**a**) Average latency for the 'create' operation. (**b**) Average latency for the 'retrieve' operation. (**c**) Average latency for the 'update' operation. (**d**) Average latency for the 'delete' operation.

5. Related Work

The increasing focus on cybersecurity in smart grids, due to the criticality of the energy infrastructure, has led to diverse approaches in recent research. Bouramdane et al. [19] provide a comprehensive exploration of cyberattacks in smart grids, emphasizing the importance of multi-criteria decision making for cybersecurity options. They propose an analytical hierarchy process to evaluate cybersecurity measures, including those incorporating artificial intelligence, which is crucial for enhancing the resilience of smart grids against cyber threats. Ref. [20] presents a survey on the application of deep learning for cybersecurity in smart grids. The authors review the latest advancements in deep learning techniques and their application in detecting and addressing cyber threats in smart grids, providing a framework for enhancing cyberattack detection. Manbachi and Hammami [21] introduce the Virtualized Experiential Learning Platform (VELP) for smart grids and operational technology cybersecurity. They demonstrate how virtualization technologies, including digital twins and AI, can be utilized in hands-on training programs, which is essential for the development of a skilled workforce in the era of Industry 4.0. Moreover, Cali et al. [22] highlight the role of digital twins in transforming energy systems and smart cities, noting their potential to improve cybersecurity, efficiency, sustainability, and reliability. Habib et al. [23] propose a digital twin approach for smart meters, focusing on solving implementation challenges and enhancing smart grid network security. Their work evaluates the communication performance of a digital twin system in a smart grid application, demonstrating its effectiveness in real-time scenarios. He et al. [24] explore a digital twin of the Energy Internet of Things (EIoT) and introduce a data-driven situation awareness paradigm to address the challenges of traditional model-based situation awareness in smart cities. In the broader context of digitalization, Barkakoti and Joshi [25] discuss the impact of digital twins on sustainable agriculture, which, while not directly related to smart grids, underscores the versatility and potential of digital twin technology in various sectors. Ramos et al. [26] present a smart water grid with a digital twin for water infrastructure management, demonstrating the utility of digital twins in monitoring and managing system efficiency. Szczepaniuk and Szczepaniuk [27] analyze the use of artificial intelligence algorithms in the energy sector, including their application in cybersecurity, smart grid management, and energy saving. Their work identifies open research challenges for the practical application of AI in critical energy domains. Lastly, Akkad et al. [28] develop an information security model for IoT-enabled smart grids, addressing the need for proactive security measures in the bi-directional data flow within smart grids. Our work distinctively combines digital twin technology with a data space to augment smart grid resilience, focusing on open-source technologies and grounding our framework in the Common Information Model (CIM), a prevalent standard in industrial settings. The adoption of a digital twin enables the correlation of information from diverse sources, which has been proven to be an effective approach for improving IDS performance [29].

The architecture's reliance on open-source technology not only guarantees cost efficiency but also encourages updates and improvements driven by the community, ensuring that the system remains flexible and current with evolving trends and practices in cybersecurity. This approach, leveraging current technologies with specific capabilities, provides a general yet practical architecture, distinguishing our work from the existing literature in the field.

6. Conclusions and Future Work

In our pursuit to safeguard critical infrastructure against cyber threats, the architecture we have proposed stands as a testament to the efficacy of standardized data models aligned with the data space view. This alignment is not merely a technical detail but a strategic approach that ensures interoperability and scalability across various systems and operators within the energy sector. Our architecture serves as a step towards the creation of a suite of services specifically designed to bolster the cybersecurity position of Electric Power and Energy System (EPES) operators. By adhering to standardized data models, we facilitate a seamless integration of diverse systems and enable a unified response to cyber threats. This standardization is crucial for EPES operators as it lays the groundwork for a collaborative defence strategy, enhancing the overall resilience of critical infrastructures. The implementation of our architecture within the smart grid has underscored the importance of real-time monitoring and rapid response capabilities. It has proven to be a viable solution that not only detects and communicates anomalies but also adheres to the principles of a data space, where information sharing and system interoperability are key. The open-source nature of the proposed architecture not only ensures cost-effectiveness but also fosters community-driven enhancements and updates, keeping the system agile and up-to-date with the latest cybersecurity trends and practices.

Looking ahead, our commitment is to broaden the reach and enhance the functionality of our architectural framework. In fact, system monitoring presents some drawbacks. Without recognising the state of the system, the architecture might be limited in its ability to conduct comprehensive behavioural analyses. State-based monitoring enables the identification of anomalous behaviours that could indicate a cybersecurity threat. Moreover, incident management is also needed in the approach and must be integrated into our proposal. Hence, our future work will focus on the creation of an extensive service catalogue, which will encompass a range of cybersecurity services. This catalogue will serve as a resource for Electric Power and Energy System (EPES) operators, offering them ready-to-implement solutions that adhere to a uniform standard for threat identification and mitigation. Simultaneously, we intend to cultivate a cooperative environment dedicated to cybersecurity. This will be a space where EPES operators can exchange knowledge, operational best practices, and actionable threat intelligence. By capitalizing on the data space model, this collaborative platform will significantly contribute to the collective fortification of security measures across the sector.

Author Contributions: Conceptualization, L.C., R.N. and L.R.; methodology, R.N.; software, A.P.; validation, L.C., R.N. and A.P.; data curation, A.P.; writing—original draft preparation, R.N. and A.P.; writing—review and editing, L.C. and L.R.; supervision, L.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101020560 CyberSEAS. The content of this publication reflects the opinion of its authors and does not, in any way, represent the opinions of the European Union. The European Commission is not responsible for any use that may be made of the information that this publication contains.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of these data. Data was obtained from Operato and are available from the authors with the permission of Operato.

Acknowledgments: The authors want to sincerely acknowledge Andrej Souvent from Operato (Slovenia) who supported our research by providing information about the real-world scenario used for the validation of the proposal.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

API	Application Program Interface
CIM	Common Information Model
DA	Data Acquisition
DT	Digital Twin
EPES	Electrical Power Energy Systems
GDPR	General Data Protection Regulation
IDMEF	Intrusion Detection Message Exchange Format
IDSA	International Data Spaces Association
IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
JSON	JavaScript Object Notation
LD	Linked Data
NGSI	Next-Generation Service Interface
NIS	Network and Information Systems
RDF	Resource Description Framework
SIEM	Security Information and Event Management
UML	Unified Modeling Language
XML	eXtensible Markup Language

References

1. European Parliament; Council of the European Union. *Directive (EU) 2022/2555 of the European Parliament and of the Council on Measures for a High Common Level of Cybersecurity across the Union*; Official Journal of the European Union: Aberdeen, UK, 2022. Available online: https://eur-lex.europa.eu/eli/dir/2022/2555 (accessed on 7 November 2023).
2. Brucherseifer, E.; Winter, H.; Mentges, A.; Mühlhäuser, M.; Hellmann, M. Digital Twin conceptual framework for improving critical infrastructure resilience. *at-Automatisierungstechnik* **2021**, *69*, 1062–1080. [CrossRef]
3. Masi, M.; Sellitto, G.P.; Aranha, H.; Pavleska, T. Securing critical infrastructures with a cybersecurity digital twin. *Softw. Syst. Model.* **2023**, *22*, 689–707. [CrossRef]
4. Grasselli, C.; Melis, A.; Rinieri, L.; Berardi, D.; Gori, G.; Al Sadi, A. An industrial network digital twin for enhanced security of cyber-physical systems. In Proceedings of the 2022 International Symposium on Networks, Computers and Communications (ISNCC), Shenzhen, China, 19–22 July 2022; pp. 1–7.
5. European Commission. A European Strategy for Data. European Commission—Digital Strategy. 2023. Available online: https://digital-strategy.ec.europa.eu/en/policies/strategy-data (accessed on 7 November 2023).
6. Coppolino, L.; D'Antonio, S.; Mazzeo, G.; Romano, L.; Sgaglione, L. How to protect public administration from cybersecurity threats: The COMPACT project. In Proceedings of the 2018 32nd International Conference on Advanced Information Networking and Applications Workshops (WAINA), Krakow, Poland, 16–18 May 2018; pp. 573–578.
7. FIWARE Foundation e.V. FIWARE—Open APIs for Open Minds. 2023. Available online: https://www.fiware.org/ (accessed on 7 November 2023).
8. Uslar, M.; Specht, M.; Rohjans, S.; Trefke, J.; González, J.M. *The Common Information Model CIM: IEC 61968/61970 and 62325-A Practical Introduction to the CIM*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.
9. Coppolino, L.; Nardone, R.; Petruolo, A.; Romano, L.; Souvent, A. Exploiting Digital Twin technology for Cybersecurity Monitoring in Smart Grids. In Proceedings of the 18th International Conference on Availability, Reliability and Security, Benevento, Italy, 29 August–1 September 2023. [CrossRef]
10. Tao, F.; Qi, Q.; Wang, L.; Nee, A. Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: Correlation and comparison. *Engineering* **2019**, *5*, 653–661. [CrossRef]
11. Latif, K.; Sharafat, A.; Seo, J. Digital Twin-Driven Framework for TBM Performance Prediction, Visualization, and Monitoring through Machine Learning. *Appl. Sci.* **2023**, *13*, 11435. [CrossRef]
12. Velazquez, A.; Martell, F.; Sanchez, I.Y.; Paredes, C.A. Cyberphysical System Modeled with Complex Networks and Hybrid Automata to Diagnose Multiple and Concurrent Faults in Manufacturing Systems. *Appl. Sci.* **2023**, *13*, 10603. [CrossRef]
13. Object Management Group. *Unified Modeling Language*; Object Management Group: Needham, MA, USA, 2001; Volume 105.

14. Bernardi, S.; Gentile, U.; Marrone, S.; Merseguer, J.; Nardone, R. Security modelling and formal verification of survivability properties: Application to cyber–physical systems. *J. Syst. Softw.* **2021**, *171*, 110746. [CrossRef]
15. Van Ruijven, B.J.; De Cian, E.; Sue Wing, I. Amplification of future energy demand growth due to climate change. *Nat. Commun.* **2019**, *10*, 2762. [CrossRef] [PubMed]
16. Olabi, A.; Abdelkareem, M.A. Renewable energy and climate change. *Renew. Sustain. Energy Rev.* **2022**, *158*, 112111. [CrossRef]
17. Kyriakou, D.G.; Kanellos, F.D. Sustainable Operation of Active Distribution Networks. *Appl. Sci.* **2023**, *13*, 3115. [CrossRef]
18. Debar, H.; Curry, D.; Feinstein, B. *RFC 4765: The Intrusion Detection Message Exchange Format (IDMEF)*; RFC Editor: Marina del Rey, CA, USA, 2007.
19. Bouramdane, A.A. Cyberattacks in Smart Grids: Challenges and Solving the Multi-Criteria Decision-Making for Cybersecurity Options, Including Ones That Incorporate Artificial Intelligence, Using an Analytical Hierarchy Process. *J. Cybersecur. Priv.* **2023**, *3*, 31. [CrossRef]
20. Ruan, J.; Liang, G.; Zhao, J.; Zhao, H.; Qiu, J.; Wen, F.; Dong, Z.Y. Deep learning for cybersecurity in smart grids: Review and perspectives. *Energy Convers. Econ.* **2023**, *4*, 233–251. [CrossRef]
21. Manbachi, M.; Hammami, M. Virtualized Experiential Learning Platform (VELP) for Smart Grids and Operational Technology Cybersecurity. In Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management, Piscataway, NJ, USA, 14–16 December 2022. [CrossRef]
22. Cali, U.; Dimd, B.; Hajialigol, P.; Moazami, A.; Gourisetti, S.; Lobaccaro, G.; Aghaei, M. Digital Twins: Shaping the Future of Energy Systems and Smart Cities through Cybersecurity, Efficiency, and Sustainability. In Proceedings of the 2023 International Conference on Future Energy Solutions (FES), Vaasa, Finland, 12–14 June 2023. [CrossRef]
23. Habib, M.Q.; Shoukat, M.U.; Irfan, M.; Zubair, M.; Ahmed, S.; Raza, M.; Ali, T.; Sarwar, A. Smart Meter Development Using Digital Twin Technology for Green Energy Distribution Optimization. *Eur. J. Theor. Appl. Sci.* **2023**, *1*, 181–190. [CrossRef] [PubMed]
24. He, X.; Ai, Q.; Wang, J.; Tao, F.; Pan, B.; Qiu, R.C.; Yang, B. Situation Awareness of Energy Internet of Things in Smart City Based on Digital Twin: From Digitization to Informatization. *IEEE Internet Things J.* **2023**, *10*, 7439–7458. [CrossRef]
25. Barkakoti, C.; Joshi, S. Advancement of Digital Twin in Irrigation and Smart Farming. In Proceedings of the 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 23–25 March 2023. [CrossRef]
26. Ramos, H.; Kuriqi, A.; Besharat, M.; Creaco, E.; Tasca, E.; Coronado-Hernández, O.E.; Pienika, R.; Iglesias-Rey, P. Smart Water Grids and Digital Twin for the Management of System Efficiency in Water Distribution Networks. *Water* **2023**, *15*, 1129. [CrossRef]
27. Szczepaniuk, H.; Szczepaniuk, E. Applications of Artificial Intelligence Algorithms in the Energy Sector. *Energies* **2023**, *16*, 347. [CrossRef]
28. Akkad, A.; Wills, G.; Rezazadeh, A. An Information Security Model for an IoT-enabled Smart Grid in the Saudi energy sector. In Proceedings of the 2022 Saudi Arabia Smart Grid (SASG), Riyadh, Saudi Arabia, 12–14 December 2022. [CrossRef]
29. Coppolino, L.; Romano, L.; D'Antonio, S.; Esposito, M. Exploiting diversity and correlation to improve the performance of intrusion detection systems. In Proceedings of the 2009 International Conference on Network and Service Security, Paris, France, 24–26 June 2009; pp. 1–5.