This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3580055

-[1]Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2022.Doi Number

Digital Twins for Power Systems: Review of Current Practices, Requirements, Enabling Technologies, Data Federation and Challenges

May Myat Thwe (Student Member, IEEE), Alexandru Ştefanov (Member, IEEE), Vetrivel Subramaniam Rajkumar (Student Member, IEEE) and Peter Palensky (Senior Member, IEEE)

Department of Electrical Sustainable Energy, Delft University of Technology, Mekelweg 4, 2628 CD, Delft, The Netherlands

Corresponding author: May Myat Thwe (e-mail: mmyatthwe@tudelft.nl).

This work has received funding from the European Commission in Horizon Europe TwinEU project under the Grant Agreement 101136119.

ABSTRACT The Digital Twins (DT) have emerged as the technology that provides capabilities to simulate and analyze cyber-physical systems' behaviors using digital replicas. This is achieved through high-fidelity digital models, bi-directional communication and (near) real-time data exchange between physical real-world systems and DTs. Despite its capabilities of facilitating real-time monitoring, optimization, and predicting system performance, effectively leveraging DT for power system applications requires integrating data from heterogeneous sources and addressing various data related aspects. These include data modeling, exchange and interoperability. One promising concept to address these aspects is that of data federation which promotes interoperability, allowing DTs to operate autonomously, yet interact seamlessly. While various studies in literature have addressed DT applications, technologies, and challenges, a comprehensive review on the data federation aspects within power systems still needs to be investigated. This research seeks to bridge this gap by providing an in-depth review of DT practices in academia and industry, functional and non-functional requirements, and enabling technologies, with emphasis on data federation. Its role in enhancing system-wide interoperability in the power system, along with associated challenges are summarized and discussed.

INDEX TERMS Digital Twin, Power Systems, Data Federation, Digital Technologies, Data Interoperability

I. INTRODUCTION

Digitalization is revolutionizing the energy sector, enhancing efficiency, and offering low-carbon energy solutions. However, the energy transition in modern power systems comes with its challenges. The growth of intermittent and decentralized renewable sources of energy and the interconnection amongst autonomous and distributed constituent systems in the power system have increased the complexity and uncertainty of power system operations [1]. This necessitates a solution that can facilitate and coordinate the interplay and collaboration between constituent systems and analyse and observe uncertain and dynamic power system behaviors. The Digital Twin (DT) concept is an approach that capture the intricate behaviours of physical systems through high-fidelity analytical models [1]. A DT is a unique combination of the physical and digital worlds. The physical world constitutes the operational technologies such as sensors, and actuators, while the virtual digital space hosts digital replicas of these physical assets. This can simulate different conditions, and configurations and take decisions regarding the physical space via (live) data and information flows between them [2], [3].

DTs allows for integrating data across power system components, for example, to predict system behaviours, and simulate interactions among these constituent systems to anticipate potential events. These capabilities are crucial in dynamic and uncertain environments, enabling real-time monitoring, data-driven optimization, and analysis, and ensuring pre-emptive measures for power system reliability and resilience. The DT has drawn significant attention from researchers in industry and academia. Its widespread adoption across various industries, including aviation, manufacturing and the energy domain shows its credibility and potential. DT plays a vital role in a variety of applications and use cases of

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3580055

future power systems, including power generation and distribution, energy storage, project planning, microgrids, electric vehicles, and renewable energy generation systems [4], [5]. The power systems digital transformation through DTs represents a significant leap forward in optimizing, managing, and securing the power systems [6], [2].

A. MOTIVATION

While DT brings various capabilities, it is necessary to effectively utilize its full potential in power systems to address the challenges such as modelling and management of data, computational requirements, and interoperability of complex interconnected and heterogeneous power systems [7],[8]. DT introduces requirements, such as the need for interoperable data and data model exchanges to enhance system-wide interoperability. Thus, a robust foundation in data federation becomes a prerequisite and the core for developing an effective DT. Data federation is an approach that virtually integrates and manages data from heterogenous sources into a common federated query engine. It integrates data from diverse sources like relational databases, structured files, through a unified schema and allows for streamlining access to distributed data as a single source.

The federated concept has been applied for a few reasons, including enhanced system interoperability, collaboration, and synchronization, but also for comprehensive and real-time decision-making while addressing the integration challenges of collaborative DTs [9]. The federated concept allows for a scalable and flexible architecture where individual DTs, each modelling specific assets or processes within the power grid, operate autonomously, yet can interact seamlessly with one another. This interconnectedness is essential for simulating complex, system-wide scenarios, ranging from monitoring to emergency response strategies, with high fidelity and real-time data exchange [2]. Through these interactions, DTs can collaborate, share data and insights that enhance decision-making processes, and optimize power system performance.

B. RELATED SURVEYS AND CONTRIBUTIONS

The literature in recent years has investigated and reviewed different dimensions of the DT, including challenges, requirements, supporting technologies and applications in general. However, only a few covered the data federation aspects of DTs in power systems. Sifat et al. [8] underscored that the design of DT should consider Cyber-Physical Systems (CPS) requirements such as security, scalability, and confidentiality while also reviewing the challenges in data communication, protocols, power grid integration, and cybersecurity. Similarly, Chen et al. [4] studied data privacy, cybersecurity, and data and model fusion challenges and how DT can enhance the control, design, and maintenance of Power Electronics Enhanced Cyber-Physical Systems (PEECSs). In [10], the survey was provided by exploring DT applications and functions and discussing DT challenges in power systems. Another survey in [11] emphasized the non-functional requirements for DT operations and presented different use case applications of DT for electrical energy.

In addition, Jeong et al. [12] defined a DT as an intelligent technology platform that synchronizes physical entities e.g., spaces, objects, processes and systems. Palensky et al. [1] examined DT applications and use cases in future power systems and their supportive functions. Brosinsky et al. [2] thoroughly investigated the importance of data federation and its required components to prevent data silos in power system DTs, along with discussing the non-functional requirements DTs should meet. The authors of [13] reviewed DT challenges, such as data ownership, governance, security, and fidelity, alongside social and ethical concerns and provided insights on the role of DT in smart factories, cities, and buildings. In addition, Fuller et al. [14] reviewed DT and data analytic challenges, detailing the technologies and functional blocks necessary for digital twinning and its application in healthcare, manufacturing, and smart cities. Zhang et al. [15] focused on the data requirements, principles, and technologies that can satisfy these requirements.

Existing studies have discussed different aspects, from DTs' applications and operational requirements to the significance of data integration. However, a comprehensive review is needed that provides the functional, non-functional, and technology requirements, case studies on DT implementation using digital technologies, current practices of DT in both academia and industry, in addition to data federation requirements, and challenges of power systems DT in detail. This gap identified in the existing literature forms the basis for our review paper's contributions. This paper comprehensively investigates several critical aspects related to DT in the energy domain.

- Firstly, we introduce an extensive overview of different DT definitions and concepts found in previous studies. Then we examine its applications across different areas within the power system context. This overview identifies trends in the existing research, providing a foundation to assimilate the state-of-the-art DT technology.
- Secondly, we investigate the specific requirements necessary for implementing DTs within power systems. This involves a careful examination of the technical and functional needs that must be met to successfully deploy DT in this context. We also explore the enabling digital technologies that support digital twinning, such as Internet of Things (IoT), Artificial Intelligence (AI), and Machine Learning (ML), and how these digital technologies integrate with power systems DT to enhance performance, reliability, and efficiency. These findings can provide implications on current trends and advancements in those technologies that are advancing DT capabilities and further research and development.
- Finally, we discuss the critical aspects of data federation within power system DTs. This includes a

comprehensive investigation of the various data types, attributes, and requirements necessary for effective digital twinning. We outline the principles and components essential for data modelling, which ensure that the interoperability standards are met. Moreover, we identify and discuss the challenges associated with data federation, including data integration, governance, management, standardization, and security issues, and propose potential solutions to overcome these encumbrances.

II. LITERATURE REVIEW

A. RESEARCH OBJECTIVES AND PAPER STRUCTURE

A thorough review of the literature was conducted to identify the state-of-the-art research. The review's objectives were defined to guide the review process and identify relevant studies and findings.

Research Objectives

As DTs gain increasing relevance in various fields, reviewing how DT is defined in the literature is imperative. Understanding its conceptual underpinnings and potential applications is also crucial, especially as DT technology has been broadly used in the energy sector. Therefore, examining current academic and industry practices documented in the literature is necessary to specify trends, use cases, and potential. Furthermore, identifying functional and non-functional requirements is vital to develop effective DTs for power systems, as they are essential for creating robust and efficient models. In addition, since digital technologies can unlock new capabilities for DTs, exploring how these technologies enable enhanced performance and innovation is essential. While data integration is an important element for the success of DTs in power systems, identifying the types of data and their attributes required for effective data federation is of utmost importance. To ensure the seamless operation of DTs, it is also necessary to define specific data requirements, principles, and components. Given the complexity of integrating diverse systems, understanding how to ensure interoperability for data federation in heterogeneous power system DTs is indispensable for achieving cohesive and functional integration. Finally, because DT data federation in power systems faces various challenges, it is crucial to identify and address these barriers to facilitate smooth implementation and operation.

Consequently, in this study, we limit our focus to the topics below:

1. Definition, concepts and characteristics of DTs based on the literature,
2. Current state-of-the-art practices and applications of DTs in the energy industry and academia,
3. Functional and non-functional requirements necessary to develop DTs of power systems,
4. Role of digital technologies for power system DTs and case studies,
5. Data requirements, principles, and components for data federation of DTs,

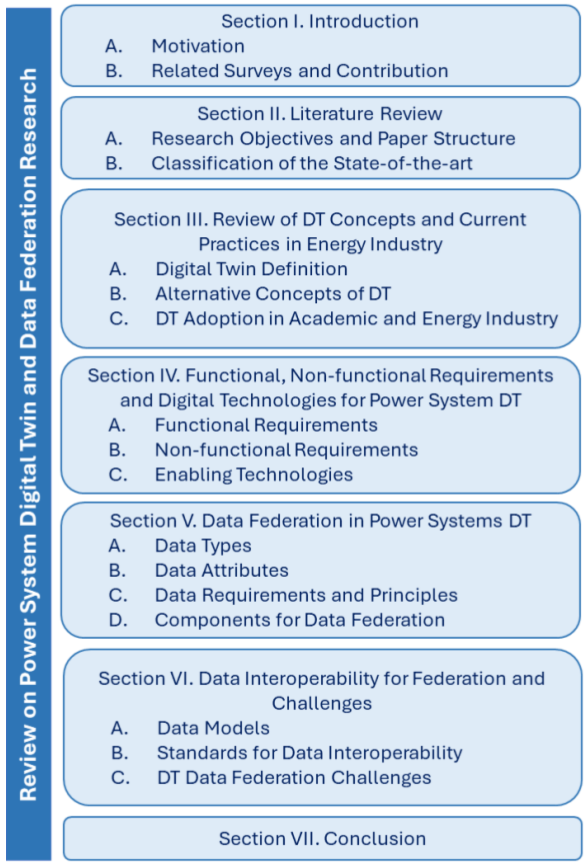

FIGURE 1. **Summary of review paper structure and coverage.**

6. Data types and attributes for data federation of power systems DTs,
7. Interoperability standards, its adoption in real-world applications and associated challenges for power system DT data federation.

B. CLASSIFICATION OF THE STATE-OF-THE-ART

This advanced review paper includes articles published in three digital libraries, i.e., IEEE Xplore, Science Direct, and Digital Twin. We limited the article category to journals, magazines, and book chapters, and the period from 2002, when the concept of DT concept was first proposed, until the first quarter of 2024 (Q1 2024). To find relevant information regarding the DT adoption in the energy industry, we searched on Google News using the search queries. This survey uses relevant Text Search (TS) queries based on the search objectives to find relevant articles from the digital libraries and Google News. Table 1 summarizes the text query used for this survey.

This comprehensive review is categorized into three sections as illustrated in Figure 1. Furthermore, to comprehensively understand the current research landscape, we analysed a research gap as presented in Table II. It compares existing survey and review articles relevant to DTs in the power system domain across five categories. These categories were selected based on their foundational role in developing robust, interoperable, and scalable DTs for power systems.

- Applications & Use Cases: This dimension evaluates how comprehensively each study addresses DT applications in power systems. Articles such as [1], [10] and [13] provide a high-level discussion of potential applications in power systems.
- Functional & Non-functional Requirements: DT systems must satisfy various requirements to be viable in power systems, such as interoperability and cybersecurity. This category assesses the degree to which prior works define or formalize such requirements.
- Enabling Technologies are discussed in most papers to a moderate extent (e.g., [4], [13], [14]), mentioning IoT, AI, or cloud computing. However, the interconnection between these technologies and integration with DT implementations in power systems remains underrepresented in prior work. We evaluate how these technologies are addressed and contextualized within the power system landscape.
- Data Federation Requirements: As modern power systems become increasingly decentralized and data-intensive, federated architectures are essential for unifying diverse data sources. The most significant gap appears in this aspect. While [2] and [14] studied these topics, their focus remains high-level and not specific to power system context. Data principles were explored in [15], but there was limited discussion on federation frameworks and standardization protocols.
- Data Federation Challenges: Effective implementation of data federation in DTs can be constrained by interoperability limitations, lack of standardization, and governance issues. We examine which issues are addressed in existing studies and to what extent.

Table II shows that related prior studies offer partial or minimal coverage of these last two categories, particularly data federation requirements, and challenges, which are crucial for scalable and interoperable DT systems. Our study, in contrast, provides high and in-depth coverage across all five dimensions, with a particular focus on the underrepresented area of data federation. It discusses data types and modelling attributes relevant to power system DTs and highlights practical implementation challenges of industrial interoperability standards, data governance issues, and security considerations.

III. REVIEW OF DIGITAL TWIN CONCEPTS AND CURRENT PRACTICES IN ENERGY INDUSTRY

This section presents the different DT definitions in the literature and other similar concepts surrounding DT, along with DT adoptions in the energy industry and academic research.

A. DIGITAL TWIN DEFINITION

In 2002, Grieves first proposed the term "digital twin" in to describe a new concept in product life cycle management. Nowadays, much of the literature has used distinct definitions of DTs [16]. A DT is a virtual replica of a physical entity. It includes the environment supporting it, which requires a standardized architecture, data format, communication protocol and end-to-end connectivity, facilitating interaction between the virtual replica and physical counterpart [17]. Grieves and Vickers define a DT using the terms twin, prototype, instance [16], [18]. The twin is the virtual representation of the real subject that accurately depicts it. The prototype is the description that contains sufficient details to reproduce the twin. The subject itself is an instance. Jafari et al. [10] denote the DT concept as a digital replica of a physical system that can reflect its physical behaviour through interaction in real-time and bidirectional data.

B. ALTERNATIVE CONCEPTS OF DIGITAL TWIN

There are various misconceptions about DTs, and clarifying these misconceptions is essential for understanding and leveraging DT technology. In that regard, Fuller et al. [14] addressed various concepts surrounding DTs, differentiating three concepts: DT, digital shadow, and digital model.

"Digital Model" is a static, digital depiction of an existing or conceptual physical entity. It lacks real-time data exchange with its physical model—for example, product design and buildings design. When automated data synchronization is absent, modifications made to the physical entity are not reflected in the digital model and vice versa. This concept is synonymous with the "modelling and simulation" concept, described by Wagner et al. [7], which provides a framework for exploring object behavior without direct interaction. "Modelling and Simulation" includes digitally replicating a physical asset to study its behavior without directly experimenting with the actual object. Developing a model is to capture essential features of the physical entity for targeted analysis and study the specific scenario's behavior and/or performance of the actual object through relevant attributes and relationships. In contrast to DT, models are created to emulate or simulate the object but focus on relevant aspects for the investigations at hand.

"Snapshot Twin" is another term utilized for simulations, and it can capture and isolate data at a particular moment for subsequent analysis. It is helpful in "what-if" scenarios, real-time simulations, and virtual reality applications. It also maintains all data regard to the physical entity within the DT database, ensuring that the snapshot and its results remain integrated within the DT ecosystem. The "Digital Shadow" concept represents a digital illustration of a real physical object characterized by a uni-directional data flow from the physical to the digital model. Adjustments made in the physical object updates only the digital object. Unlike Snapshot Twin and Digital Shadow, "Digital Twin" provides a dynamic, bidirectional data exchange and synchronization between a physical and digital object. Modifications and adjustments in the actual physical space immediately update the digital version, and vice versa.

Similarly, "Linked DT" is a comprehensive digital representation that is technically feasible and applicable to various experimental scenarios. Simulations in this context are conducted separately from the physical entity, focusing on how the object might respond to specific conditions or inputs. Facilitating direct and two-way communication as well as real-time updates will allow for close alignment between the DT and its physical counterpart. In essence, a DT in the energy domain is a virtual representation of the actual physical entity (such as power grid assets, power system operation processes) that can mirror its state, behaviour and performance in the virtual space. Physical entity and DT convergence occurs as often as needed, with an appropriate rate of synchronization and bi-directional data communication.

TABLE I
TEXT SEARCH QUERIES TO INVESTIGATE RELATED LITERATURE ON POWER SYSTEM DIGITAL TWIN AND DATA FEDERATION IN DIGITAL TWIN AND ITS CHALLENGES RESEARCH

Objectives	Text Search Queries
Part one: Search relevant articles on digital twin in power system context	TS1: `((survey OR review) AND (digital twin) AND (power system OR smart grid))` TS2: `((power system digital twin) AND (application OR use case OR case study))` TS3: `((digital twin) AND (energy industry OR energy companies))`
Part two: Search relevant articles related to DT functional, non-functional requirements and digital technologies.	TS4: `((functions OR requirements) AND (digital twin))` TS5: `((technologies) AND (digital twin))`
Part three: Search relevant articles related to data federation in digital twin and challenges	TS6: `((data federation OR data modelling) AND (digital twin) AND (power system OR smart grid))` TS7: `((digital twin) AND (challenges) AND (power system))`

TABLE II
COMPARISON OF RELATED WORK ON POWER SYSTEM DIGITAL TWIN SURVEY AND REVIEW

Reference	Applications & Use Cases	Functional & Non-Functional Requirements	Enabling Technologies	Data Federation Requirements	Data Federation Challenges	Remarks
[1]	H	L	M	L	L	Examined the DT concept evolution, features and DT applications in future power systems and functions.
[2]	L	H	L	H	M	Discussed the concept of DT in the energy management system and examined data federation aspects and requirements to construct a comprehensive model of a digitalised power system.
[4]	M	L	M	L	L	Summarized key technologies of DT, applications, and use cases of DT for design, control, and maintenance.
[8]	L	M	H	L	L	Studied requirements and DT grid challenges and opportunities and proposed theoretical framework of the DT grid.
[10]	L	H	M	L	M	Explored DT applications and functions and discussed DT challenges in power systems.
[11]	M	H	M	L	L	Discussed the operational requirements of DT, classified security threats in DT paradigm and presented different use cases for DT applications in different industries.
[13]	H	L	H	L	M	Provided a detailed investigation on advantages and use cases of technologies in different domains and researched challenges for future developments in DTs.

[14]	H	M	H	L	H	Reviewed functional blocks and enabling technologies for DT, discussed the data related challenges and provided an overview of DT adoption in smart cities, manufacturing, and healthcare.
[15]	L	L	H	M	L	Provided a thorough review of data requirements and principles of DT that need to be considered when integrating data.
This paper	H	H	H	H	H	Provided a detailed review of DT applications in academia and industry, identifying functional, non-functional requirements and technologies that enable digital twinning in power system, as well as data federation aspects of DT focusing on data interoperability and challenges.

H: high and in-depth coverage of the subject M: partial coverage L: brief or no coverage

As discussed above, one of the fundamental characteristics of a DT is its ability to collect and process data from its physical counterpart and send feedback, insights, and outcomes back to them. The integration of DT results into physical systems relies on robust and reliable communication infrastructures. DTs and physical systems use protocols such as Transmission Control Protocol/Internet Protocol (TCP/IP) as the foundation for real-time communication and industry-specific standards—such as International Electrotechnical Commission (IEC) 61850 in power systems or OPC UA in industrial automation, depending on specialized data handling needs. This enables the transmission of measurements and real-time operating data from the physical system flow to the DT for processing and analysis. DT-generated insights, such as control and command parameters and operational strategies, are sent into the physical system at appropriate synchronization rates to ensure consistent system states. Different systems can use different protocols for integration based on the specific applications and use cases. Shen et al. [19] provide an example of DT integration with a physical system using real-time data. The authors suggested a testbed for establishing a power system DT (PSDT) that uses TCP/IP communication protocols and synchronization between physical and DTs. They demonstrated the DT's practical applicability using data and scenario generation scenarios, online fault identification, measurement upscaling, and expansion. While DT and physical systems often share a common communication infrastructure, achieving seamless integration relies on particular system requirements such as data heterogeneity, real-time processing demands, resilience to communication disruptions, and cybersecurity risks.

C. DIGITAL TWIN ADOPTION IN ACADEMIA AND INDUSTRY

1) ACADEMIC PRACTICES

Much research in recent years has discussed how DTs have been applied in the energy domain. Figure 2 presents different areas where DT has been utilized in the energy domain. One of the most widely adopted applications is to improve the reliability and performance of power converters and other critical equipment and enable real-time monitoring and predictive analytics for addressing potential failures [20]. Microgrids have applied DT technology for optimized integration and management, ensuring a stable and efficient operation in complex energy distribution scenarios [7]. In addition, DTs facilitate the development of more efficient and reliable transportation solutions, including Electric Vehicles (EV) and drive systems, which align with energy storage systems for seamless energy flow and storage capabilities [21]. Renewable energy sources, including solar panels and wind turbines, can also benefit from DT to maximize output and effectively integrate renewable resources into the power grid [22].

Furthermore, DTs extend their applications to broader aspects of energy management, including forecasting to optimize energy usage in smart cities, cyber-physical attack detection to secure the power grid infrastructure, and energy forecasting [23],[24], [25]. They have also contributed to fault prediction, detection, and diagnosis in Photovoltaic (PV) energy conversion units, power converters, and distributed PV systems. DT can enable precise identification of operational issues in PV units, supporting early detection of faults in power converters, facilitating detailed fault diagnosis in distributed PV setups [26], [27], [28], and fault prediction in the transmission lines [29]. Moreover, DTs are used in the design phase of energy systems including structure design, simulation platform design, and reliability design. DT allows for the simulation of various scenarios, enabling engineers to optimize designs for robustness and reliability before the actual physical implementation [21], [30]. Similarly, DTs enable more sophisticated optimization, precise prediction of state variables, and the execution of complex control tasks, thus significantly enhancing the management of control systems. Through continuous monitoring and predictive analytics, DTs provide a proactive framework for maintaining the system health, ensuring operational efficiency, and extending the lifespan of energy conversion and distribution systems [31]. As power systems increasingly integrate DERs, analysing and optimizing for sustainability becomes paramount, and approaches for DT-driven sustainability assessment and optimization of performances have been proposed [32]. As DT has the potential for the automated gathering of data directly from power generation, they can be utilized for analytics using ML to produce meaningful insights, enabling informed decision-making for reducing energy consumption, minimizing waste, and optimizing resource use.

The aforementioned power system DT solutions must be developed and tested using processes that do not interfere with regular business operations, jeopardize electrical services, or access confidential data. Nevertheless, access to physical counterparts and the use of data from actual physical power system infrastructure are frequently limited for research purposes in academic settings. In this regard, a real-world testing platform should be established by using the virtual physical twin (VPT) as a replacement for the physical counterpart in constructing the power system DT testbed. Shen et al. [19] argued that VPT is not expected to be a replica of its physical counterparts and precise emulation. They proposed the implementation of a virtual testbed in two phases. The first stage connects the VPT to the DT, while the second stage focuses on the interaction between the DT being developed and the actual system. Using the VPT in the initial phase eliminates the need to precisely replicate a physical system in a DT before specifying or creating any applications. The PSDT is deployed on a real-world power system in the second step of the workflow. That occurs after the DT and related services show robustness and satisfactory performance with the VPT interface.

2) INDUSTRY PRACTICES

The energy sector is experiencing a shift towards digitalization, driven by the adoption of DT technology, to revolutionize how energy companies and governments operate. Through various innovative collaborations and projects worldwide, application of DT in power systems can be seen as summarized in Table III. For instance, in Brazil, Companhia Energética de Minas Gerais (CEMIG) and Enline Energy Solutions have undertaken digitizing power grids using a sensor less Software as a Service (SaaS) solution. By leveraging real-time data from CEMIG's Supervisory Control and Data Acquisition (SCADA) system, system operators can optimize assets without physical sensors while reducing operational costs and carbon emissions. In addition, it also aims to enhance cybersecurity through a secure connection between operational and corporate networks, characterizing the multifaceted benefits of DTs in improving grid operations and environmental sustainability [33]. Meanwhile, the European Organization for Nuclear Research (CERN)'s collaboration with ABB in Switzerland focuses on increasing energy efficiency within its cooling and ventilation systems [34]. Deploying smart sensors and creating DTs for specific infrastructure facilitates real-time monitoring, diagnostics, maintenance, and optimization. ABB's technology can enable data-driven decision-making, significantly saving energy and reducing costs. ABB has also partnered with CORYS, a French simulation company, to advance DT modelling and simulation technologies across the energy sector and beyond, to reduce operational expenses and risks, highlighting how DT technology can enhance plant operations and maintenance strategies on a broad scale [35], [36].

Furthermore, Finland's transmission system operator, Fingrid [37], has partnered with Siemens to introduce the ELVIS digital grid model. By connecting the single source of truth (SSoT) model to asset management data and historical and real-time measurements, the DT has been used to forecast future energy consumption and develop several investment scenarios, considering different policy frameworks. Adopting the electrical DTs in Fingrid has shown that it can save time and money while improving the accuracy and consistency of network models and offering efficient digitalization of current and future business processes. American Electric Power (AEP) is another largest transmission network in the United States. It aims to coordinate network model information across multiple functional business domains and centralize its management. AEP partnered with Siemens to deploy the electrical DT solution for the network model management improvement program [37]. The solution is designed based on the open standard of the Common Information Model (CIM). It allows for efficiently maintaining, analyzing, and exchanging network data across different domains. It ultimately can reduce the time and costs associated with manual model internal and external organization coordination.

The State Grid Corporation of China (SGCC) is the leading company in China in adopting DT for power grid management [38]. Focusing on Ultra-High Voltage (UHV) and smart grid technologies, SGCC's efforts symbolize the efficiency gains and operational improvements achievable through DTs, emphasizing the technology's potential in large-scale infrastructure projects. Siemens also leverages gPROMS Digital Process twin technology, presenting how DTs are used for process optimization [39], [40]. By enabling virtual design and testing, Siemens demonstrates how DTs can significantly reduce the development time and address issues such as raw material expenses and high energy costs, marking a significant advancement in manufacturing and production processes. IBM's integration of generative AI with DT technologies further highlights the potential for

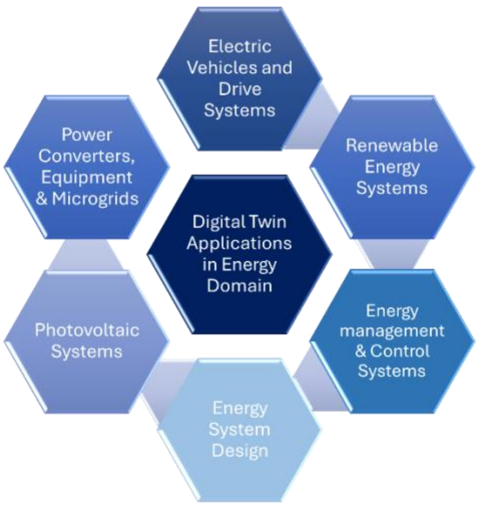

FIGURE 2. Overview of digital twin applications in energy domain.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3580055

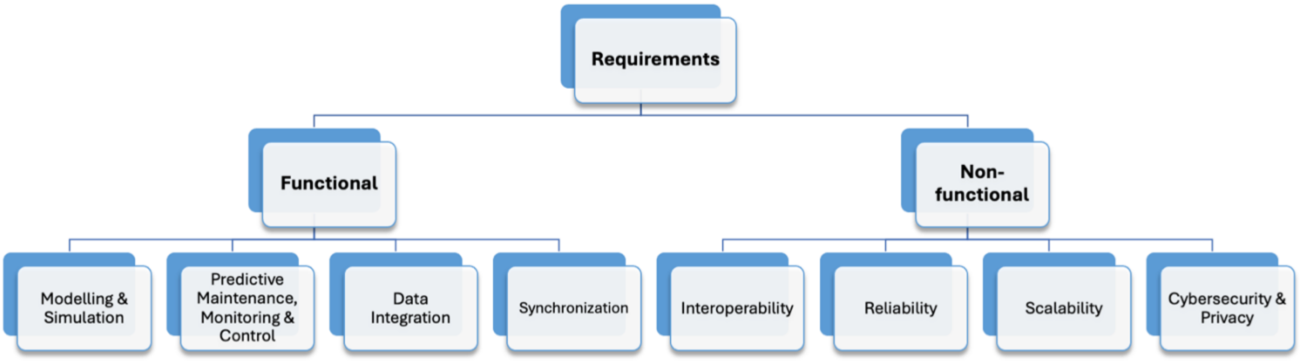

FIGURE 3. Functional and non-functional requirements for digital twin in power system.

innovation in the energy sector [41]. This approach allows for advanced asset management and operational efficiency through applications like visual insights for anomaly detection, large-scale asset performance management, and AI-powered real-time field service assistance. Through the Big Data Ecosystem and DT Platform initiatives, the government and private sector stakeholders in the United Arab Emirates (UAE) are also applying DT for creating accurate digital models of physical assets and infrastructures, to improve decision-making and risk management. By providing a dynamic 3D visualization of urban sustainability metrics, the UAE is leading the way in proactive climate change mitigation and infrastructure protection. South Korean company Techtree Innovation [42] is also contributing to the UAE's DT ecosystem, by sharing its 3D geospatial map technologies. Above worldwide industrial practices present the significant impact of DT technology in the energy sector by enhancing operational efficiency, reducing environmental impact, and fostering innovative approaches to future energy management.

IV. REQUIREMENTS AND ENABLING TECHNOLOGIES FOR POWER SYSTEM DIGITAL TWINS

When developing a DT for power systems, identifying the functional components and requirements is important to achieve a precise, and interactive representation of physical elements and processes within power systems. Functional requirements of power system define the specific behaviours or functions of the power system, detailing what the system should do or perform in terms of operations, services, and constraints. Non-functional requirements specify how the system will do it [43]. These requirements are critical for the power system design, development, and operations, ensuring it meets the needs of its users and operates efficiently and reliably. Key functional and non-functional requirements to be considered when developing DTs for the power system are depicted in Figure 3.

A. FUNCTIONAL REQUIREMENTS

1) MODELING AND SIMULATION

A DT includes models of the physical system, which are created using specific tools and are updated during the entire lifecycle. Vargas et al. [44] discussed different types of models such as submodels, complete models and synthetic models. Submodels are the decomposed parts or specific modules of a system/process. Each module is viewed as a unit with unique properties in terms of structure, behaviour, inputs, and outputs. Complete models are a collection of submodels that combine to build a larger, composite model that is accurate, sufficient, and suitable for providing services in a DT when data is received/sent to its physical counterpart. Due to the limitations of accessing real systems/processes, parameters, and data in the research community, synthetic models have been used as an alternative. It enables to investigate and assess a wide range of processes and occurrences. Although synthetic models share many of the similar features and characteristics as real systems, they can be freely shared and do not contain sensitive information. Using those models of power system assets, e.g., transformers, generators, transmission lines, Phasor Measurement Units (PMU), relays, etc., behaviours under different conditions can be simulated. DTs should have the capability to run simulations using different scenarios derived from the data from the physical system including predicting the likeliness of certain events and analyzing the impacts regarding the physical system. For example, DTs have been used to assess the stability and resilience of power systems under extreme weather events, cyber-attacks [45], [46] and renewable energy sources integration [47], [48], [49]. In addition, different simulations are performed to examine the potential causes when an anomaly is detected, or the effect on the system stability as a consequence of cyber-attacks. This requires the simulation to be (near) time to identify and mitigate the anomalies or faults. Furthermore, To optimize battery configurations as the battery runs in different environments or conditions, what-if simulations are also carried out [50]. By simulating the complex dynamics of power systems using DT, operators and/or researchers can foresee and mitigate potential issues, optimize grid performance, and prepare for future expansions or upgrades. Simulation results are crucial for understanding the effects of operational decisions and for testing the hypotheses in a virtual environment.

TABLE III
DIGITAL TWIN APPLICATIONS IN ENERGY SECTOR OF DIFFERENT COUNTRIES

Country	Company/ Institution	DT Application	Key Objectives
Brazil [113], [116]	CEMIG and EnlineEnergy Solutions	Digitization of power grids with a sensor less SaaS solution, leveraging real-time data from SCADA systems for transmission line optimization.	Reducing cost of operations and carbon emissions, enhancing cybersecurity.
Switzerland [118]	CERN and ABB	Increasing energy efficiency in cooling and ventilation systems through smart sensors and DTs for infrastructure, enabling real-time monitoring, diagnostics, optimization and maintenance.	Saving energy, reducing costs, data-driven decision-making.
France [110], [117]	ABB and CORYS	Advancement in DT modeling and simulation technologies across the energy domain and beyond.	Lowering capital and operational expenses, reducing risks, enhancing plant operations, and maintenance strategies.
China [114]	State Grid Corporation of China (SGCC)	Power grid management, focusing on UHV and smart grid technologies.	Efficiency gains, operational improvements in large-scale infrastructure projects.
Global [97]	Siemens	Leveraging gPROMS Digital Process twin technology for process optimization in manufacturing and production processes.	Reducing development time, addressing high energy costs, and raw material expenses.
Global [112]	IBM	Integration of generative AI with DT technologies for advanced asset management and operational efficiency, including applications like visual insights for anomaly detection and AI-powered real-time field service assistance.	Advanced asset management, operational efficiency improvement.
UAE & Korea [115]	Government and private sector & Techtree Innovation	Use of DT through the Big Data Ecosystem and DT Platform initiatives, focusing on creating accurate digital models of physical assets and infrastructures for improved decision-making and risk management, along with dynamic 3D visualization. Contribution to the UAE's DT ecosystem with expertise in 3D geospatial map technologies.	Proactive climate change mitigation, infrastructure protection, enhancing decision-making and risk management. Supporting sustainable energy solutions through global collaboration.
Finland	Fingrid and Siemens	ELVIS: DT for forecasting future energy consumption and developing investment scenarios by taking different policy frameworks into account.	Saving time and money while improving accuracy and consistency of network models and offering efficient digitalization of current and future business process.
USA	AEP and Siemens	DT of entire transmission network for efficiently maintaining analysis and network data exchange	Providing better coordination of network model information across multiple business functional domains and centralized management of it, while reducing cost and time

2) PREDICTIVE MAINTENANCE, MONITORING AND CONTROL

Power system DTs should provide predictive maintenance, monitoring, and control capabilities, which can be leveraged by AI and ML techniques [51], [52]. By analysing patterns in the dataset, equipment failures can be predicted before they happen, enabling operators to schedule maintenance proactively. Moreover, forecasting renewable energy production, monitoring, and managing asset health and performance, and optimization algorithms can also help minimize downtime and reduce operational costs while guiding adjustments to improve overall system stability and performance [53], [54]. Monitoring techniques can be distinguished into online and offline monitoring [55]. Online monitoring processes new data in real-time, for which delays are intolerable. In contrast, there can be delays in offline monitoring, for example, stored time series data can be utilized later for different experimentation, such as simulating events like faults or attacks and finding potential causes of events. To ensure the safety of battery physical twin, DT is used to track the discrepancy between the monitored and predicted behaviors, thereby enabling anomaly detection [50] and performing control measures as required.

3) DATA INTEGRATION

Accurate and efficient handling of data streamed from heterogeneous sources in physical power systems such as IoT devices, smart meters, and sensors requires seamless data integration and management approaches [56], [57]. It involves ensuring the aggregation and harmonization of measurement data from DERs and other components in the physical environment and updating them in real-time for analysis in the digital space. For the grid's stability and efficiency, it is paramount for DTs to have a unified view of its operational status and proactive responses to emerging system behaviors and conditions. When incorporating and exchanging data between diverse systems, such as the data integration from transmission and distribution systems, they often have disparate data models based on their operational goals and use cases. The challenge lies in enabling continuous communication and exchange of data between such systems to enhance grid management, reliability, and interoperability. For example, the CIM has been adopted to address this challenge by providing a standardized, object-oriented model for describing the electrical network and its operational data, facilitating interoperability among diverse systems [58].

4) SYNCHRONIZATION

DT should constantly be synchronized with the present state of the system, reflecting any changes that take place in the physical system [59]. However, synchronizing the changes in the physical system and in the DT is challenging [60]. While frequent synchronizations would result in higher costs and more congestion in data flow, infrequent synchronizations would cause calculation bias and imprecise decisions. It raises the question of how to enable reliable and efficient synchronization [61]. Tan et al. [62] discussed different synchronization problems including prediction update, and model update. For example, deciding when to perform a new simulation experiment to attain an updated estimate of DT performance is the goal of the prediction update. Model update entails model parameter's update, expanding detail level of model, or optimizing its parameters based on the most recent physical system observations. They studied the optimal synchronization problem as a dynamic stochastic control problem in their research. The goal is to reduce the overall misalignment costs of the prediction error resulting from the DT not being synchronized within allotted period as well as the projected synchronization cost of the DT. Jiang et al. [63] proposed optimization of Planning, Scheduling, and Execution (PSE) in precast on-site assembly using DT-enabled real-time synchronization. DT offers cyber-physical visibility and traceability, which allows for the dynamic adjustment of PSE tasks based on real-time resource status and construction progress information. Furthermore, event-based and time-based approaches can be used to synchronize the physical system and digital counterparts. For example, in [19], the DT is updated, and the system's state is evaluated every 0.5 seconds under steady-state conditions. When system variables change in the physical system, an event-based update of the DTs is initiated to reflect the observed change. Event-driven synchronization can also be performed, such as faults or breaker operations. Variables and parameters can be tailored to each application to reduce computational and communication overhead.

B. NON-FUNCTIONAL REQUIREMENTS

1) INTEROPERABILITY

The ability of two or more systems or pieces of equipment made by different vendors to communicate and use information is known as interoperability, according to the Institute for Electrical and Electronics Engineers (IEEE), and it is attained by adhering to a set of standards. For DTs to remain applicable and interoperable among the federated DT ecosystem, the design of each individual DT should satisfy interoperability requirements by standardizing data formats, communication protocols, and so on [64], [65]. There are four types of interoperability for large-scale systems: technological, syntactic, semantic, and organizational [66]. Technical interoperability emphasizes the direct data exchange between systems, requiring compatibility in technical specifications. Syntactic interoperability considers the structure and format of data exchange, ensuring that data messages are encoded in a universally understandable syntax. Semantic interoperability ensures that the data exchanged holds the same meaning across different systems. Organizational interoperability addresses the governance aspects, defining the roles and responsibilities to foster interoperability. The coalition interoperability model's layers define nine interoperability layers. It includes physical, protocols, data/object model, and information interoperability, knowledge/awareness, aligned procedures, aligned operations, harmonized strategy, and political objectives levels [67]. In the energy domain, there are existing standardizations such as IEC 61850, IEC 61970, IEC 62325, and IEC 62541 for information exchange, coordination and harmonization among different stakeholders and components. For instance, The IEC 62541-standardized Open Communication Protocol United Architecture (OPC UA) protocol has been widely used for platform-independent, service-based communication in industrial automation. It facilitates the exchange of real-time data between control devices made by various vendors.

2) RELIABILITY AND SCALABILITY

Reliability requirements need systems to consistently perform their intended functions with high quality and deliver their intended services accurately [11]. DTs monitor and simulate real-world systems' processes and environments while handling extensive amounts of real-time data generated from numerous data sources. Data is the key driver of DT. The DT of power system needs to collect high-quality data that can capture required aspects of physical power systems, validate and calibrate them. For DT to be reliable, it needs to maintain (near) real-time synchronization between the DT and physical power system to reflect the current states of the physical power system accurately. In addition, scalability ensures that a DT can expand its capabilities and continue to perform efficiently even when the scope of the DT grows [68], [69]. For instance, when extending the geographical coverage of the power grid and integrating more grid network components and data points or integrating new technologies, the design of DT should remain applicable and scale up as the system grows.

3) CYBERSECURITY AND PRIVACY

Because of the sensitive power system operations nature and the catastrophic potential of data breaches or system intrusions, cybersecurity and privacy requirements are critical [70], [71]. DT should incorporate vigorous security mechanisms to defend against cyber threats and guarantee the integrity and confidentiality of operational data [72], [73]. One of the approaches include developing the anomaly-based or signature-based intrusion detection algorithms within DT. They can monitor and differentiate the normal and anomalous network activities and identify threats by comparing them to established attack scenarios and notify an alert to protect from the cyber attacks. Furthermore, robust encryption techniques like hashing and the Advanced Encryption Standard, can been implemented to protect data used in DT from potential cyber attacks. That can help encrypt the sensitive data and protect it against unauthorized persons even if the communication between physical system and DT is being intercepted. Furthermore, blockchain has become an vital enabler for DTs, particularly for managing complexities of product lifecycle data across a diverse ecosystem of participants [74]. By leveraging a decentralized, secure, and immutable ledger, blockchain provides secure data storage, access, and sharing while maintaining the authenticity of DT data. Decentralized applications built on blockchain can support secure, owner-centric data-sharing models that uphold data integrity and confidentiality [75], [76], [77].

In addition to the cybersecurity, DT should also provide privacy when handling with sensitive information. The General Data Protection Regulation (GDPR)'s Data Protection Impact Assessment for smart grid and smart metering environment supports data controllers in establishing the rules for collecting and processing personal data. DT has been used to train ML models based on energy consumption data of household appliances to analyse electricity consumption data and predict the future energy demands. However, those consumer-specific energy data can be linked with identifying and monitoring behaviour patterns of individuals and organizations. One potential approach that have been applied is the federated learning, which is a decentralized and privacy-friendly ML [78].

C. ENABLING DIGITAL TECHNOLOGIES FOR DIGITAL TWIN

DT requires the comprehensive (near) real-time status from the physical environment. DT also needs the ability to simulate different scenarios to make smarter predictions and decisions, by collecting, analyzing, and correlating data from various physical power system components. Digital technologies are important to the implementation of DTs. They can facilitate the management, analysis, and integration of data seamlessly. They can ensure that DTs can analyze large amounts of data in real-time for more intelligent decision-making, from data generation to analysis. This section categorizes key enabling digital technologies—such as IoT, AI, blockchain, cloud computing, and edge computing—based on their specific contributions to data lifecycle stages. In addition, case studies on how those technologies can be integrated in the implementation of DTs are also discussed. Figure 4 presents the overview of digital technologies for DT implementation.

1) DATA GENERATION

Components that are part of power system infrastructure, including transformers, circuit breakers, and substations, can provide data about the power systems operational status. They are integral for different applications such as load capacity analysis, demand forecasting, identifying operational anomalies, or other dynamic stability studies. Distributed Energy Resources (DERs) including battery storage systems, wind turbines, and solar panels, can also produce data essential to manage the renewable energy to optimize grid performance. However, lack of quality and fine-grained data can hinder the development of solutions for the applications mentioned above. For example, ML based electricity demand forecasting requires substantial volume of data to train ML models and perform analysis. Despite public datasets availability to train ML models, their limited scope and size can impact obtaining highly accurate outcomes. In addition, ensuring compliance with regulations and privacy remains a challenge. Synthetic data generation approaches using Generative Adversarial Networks (GAN) is proposed to generate large-scale synthetic time-series data in smart grids to address data availability and maintain privacy, especially personally identifiable information from Advanced Metering Infrastructure (AMI) meters data [79]. Similarly, Lui et al. [80] proposed a scalable approach for creating synthetic Cyber-Physical Power Systems (CPS) topologies with realistic network characteristics. It captures real CPS networks features using graph variational autoencoders and graph neural networks (GNNs) while hiding vulnerable topological information and preserving similar features to the real networks.

2) DATA COLLECTION

Collecting new data or using existing data is a crucial step to develop an accurate DT. Integrating the IoTs within power systems is a significant shift towards more interconnected power distribution networks characterized by interconnected microgrids, and DERs. IoT devices can transform traditional

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3580055

power systems into intelligent CPS by continuously feeding large data streams from physical and cyber assets into DTs [81]. Power substations within the grid are equipped with advanced industrial IoT-based sensing and real-time monitoring systems. They can enable operators to remotely monitor substation conditions with high precision during steady and transient states [82]. These systems gather extensive data suitable for both real-time and deeper offline analysis. Bundele et al. [83] also developed a microcontroller-based phasor measurement unit (PMU) with IoT capabilities to measure voltage amplitude and frequency of power system. The prototype was designed to improve the real-time data acquisition accuracy in power systems.

3) DATA STORAGE

Large volume of data correlated both physical and digital entities, and processed data should be converted into a unified mode and stored to reuse, share and analyze. Edge computing is a distributed computing framework, which offers the computation and storage near data sources. By facilitating collection of data real-time and processing at the network's edge, near data sources, edge computing enables fast data access and processing [79]. Cloud computing is another mechanism for integrating ML and AI techniques for big data analytics. For instance, substations that are equipped with smart sensors and IoT devices can gather power flow data, voltage levels, load demands and so on in (near) real-time. Those data can be initially stored and processed at the edge locally, at the substations, allowing for immediate responses to dynamic grid conditions. Consequently, rapid adjustments such as balancing power loads or responding to system faults or disturbances can be made without the latency of data travel to the central cloud server [80], [81]. Meanwhile, cloud services provide storage and high computational resources to aggregate, store and analyse large volume of data from multiple substations for long-term analysis or forecasting. By integrating IoT data with processing techniques in the edge and cloud, a virtual representation of the power grid that dynamically updates and optimizes itself based on real-time inputs from the edge, can be created while significantly improving the overall system efficiency and performance [87].

4) DATA ANALYSIS

While data analysis is a important step in DT, AI, ML, and Deep Learning (DL) can provide meaningful analysis and insights for DTs in energy sector [88]. The consumption of renewable energy resources has been increasing, potentially surpassing traditional energy sources. Traditional methods for electrical operations such as monitoring restoration manually can lead to issues such as frequent downtimes with intermittent RES integration. Thus, they can be inadequate in addressing the complex challenges, particularly to adapt to unforeseen circumstances. The transition to smart grids for highly reliable and efficient energy management is rapidly evolving. It requires the adoption of advanced approaches to handle the big data produced by numerous components in the energy infrastructure. A more intelligent energy paradigm can be created by utilizing AI, ML and DL technologies, which integrates high intelligence into operational and supervisory decision-making. [89].

AI encompasses the broader concept of machines that can execute tasks that require human intelligence. ML, one branch of AI, represents a methodology that learns from data, improves from experience, and makes decisions. It can be classified into supervised learning (that use labeled datasets for prediction or classification), unsupervised learning (working with unlabeled data for clustering or grouping), and reinforcement learning (learning to make decisions based on feedback from performed actions) [90]. DL, sub-branch of ML, is based on multilayered neural networks, deep neural networks to train and learn from data. ML starts with data preprocessing. It includes data preparation or cleaning such as identifying and correcting errors, missing values, and duplicates. It is the foundation for data analysis, and it can have an impact on the performance and effectiveness of ML models. In addition, cleaned data needs to be transformed into suitable format, and standard while keeping the same meaning of dataset's content. Feature engineering involves scaling, normalizing, and extracting features from raw data using feature extraction techniques such as Principal Component Analysis (PCA) [91], Linear Discriminant Analysis (LDA) [92], Convolutional Neural Networks (CNN) [93], domain knowledge and so on. After preprocessing data and feature engineering, the next step is model training by dividing the dataset into training and testing sets. By selecting training algorithms or models such as decision tree [94], support vector machine [95], Long Short-Term Memory Networks [96], training data are fed to compute loss function, extract patterns and calculate the results. The performance of ML models can be assessed through testing datasets using different metrics including area under the ROC curve (AUC-ROC), accuracy, precision, recall or sensitivity, and F1 score. Integrating ML into DTs can offer significant advancements in monitoring and optimizing various systems. For instance, the health monitoring of wind turbines, anomaly detection in smart grid using PMU data, employ neural networks and genetic algorithms for the real-time power systems control and predictive health management of electric vehicle motors and photovoltaic systems [97], [98], [99], [100].

5) CASE STUDIES ON THE INTEGRATION OF DIGITAL TWIN WITH ENABLING DIGITAL TECHNOLOGIES.

The integration of DT with IoT and edge-cloud computing has been explored in [101]. The authors suggested a cloud-based DT design that facilitates, aggregates, and offers insights to help the distribution system infrastructure. A virtual representation of the networked microgrids' cyber and physical layers is an example of the proposed DT. IoT connects sensors, controllers, and actuators in energy cyber-physical systems (ECPS), enabling real-time data collection and processing. Distributed sensors provide inputs such as voltage, current, temperature, and other operational parameters from physical assets like distributed energy resources (DERs), energy storage systems, and microgrids. IoT devices utilize MQTT protocols for lightweight, efficient communication with cloud services. Amazon Web Services (AWS) hosts the virtual space for DT operations, facilitating storage, data analysis, and predictive modeling. Services like AWS Sage Maker are employed to train ML models for outage management, predictive maintenance, and optimization of grid operations. Cloud-based DTs combine data from multiple sources to create a cohesive representation of physical and cyber systems, enabling scalability for large systems. In that framework, AWS IoT Greengrass (GG) serves as the edge layer for localized data processing and immediate decision-making, reducing latency and preserving privacy. Edge systems perform preliminary data filtering and analysis to minimize communication bandwidth to the cloud. Distributed control is achieved using edge-hosted secondary controllers that regulate voltage, frequency synchronization, and power sharing among microgrids.

The proposed Hybrid DT Architecture effectively handles applications with less frequent updates, such as predictive maintenance and energy management, as well as real-time applications like contingency analysis, system restoration, and voltage regulation. The research demonstrates the feasibility and transformative potential of IoT, cloud, and edge technologies in energy system DTs. It shows the importance of future research directions, such as incorporating advanced encryption and authentication mechanisms for IoT-to-cloud communication to prevent unauthorized access, and integrating real-time intrusion detection systems to protect the edge and cloud layers from cyber threats.

Secondly, in the framework proposed by [73], DT is a virtual representation of smart grid components, simulating physical counterparts for monitoring and predictive analysis. Integrated into the Software Defined Networking (SDN) control plane, DT enables enhanced management of smart meters and grid states by storing operational behavior models and real-time analytics. IoT devices facilitate data collection and communication between grid components, such as smart meters and service providers, via protocols like MQTT. IoT devices serve as data sources and are fed into the DT. In addition, blockchain systems are used for processing and secure transmission. Blockchain provides a secure and decentralized framework for data authentication and privacy. Their approach ensured safe data sharing amongst grid components by implementing a voting-based consensus process to authenticate smart meters and establish immutable records.

In addition, Bi-GRU (Bidirectional-Gated Recurrent Unit) and a self-attention technique were also implemented to improve intrusion detection. The framework addresses vulnerabilities in open channels used by IoT devices, enhancing data integrity and mitigating risks like Man-in-the-Middle (MiTM) and Distributed Denial of Service (DDoS) attacks. The results are demonstrated using performance metrics in terms of accuracy (99.73%) and precision (97.3%), surpassing traditional methods like LSTM and GRU. It's crucial to be aware of these risks and the need for robust solutions. DL enables temporal and spatial analysis of network behaviors for real-time threat detection and response. Expanding the proposed framework to handle a larger number of IoT devices and real-time datasets and optimizing the SDN control plane for faster decision-making and lower latency in highly dynamic smart grid environments can be the extension of their work. Moreover, future research can focus on advanced security features by integrating blockchain with advanced cryptographic techniques to further secure sensitive energy data. Additionally, studies on reducing the computational overhead of DL and blockchain by exploring lightweight algorithms and edge-based processing for resource-constrained devices can also be beneficial.

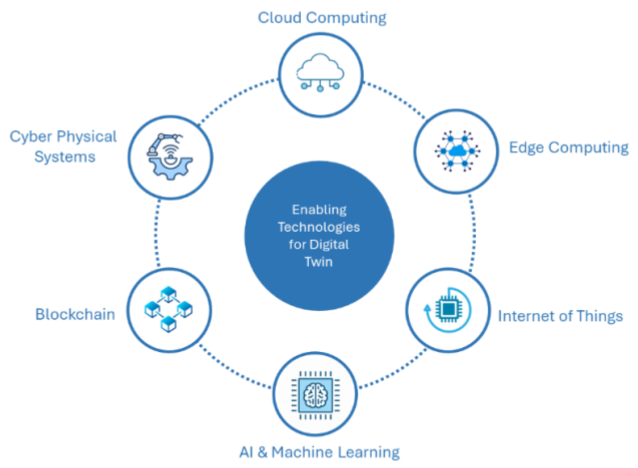

FIGURE 4. Overview of enabling digital technologies for digital twin implementation

V. DATA FEDERATION IN DIGITAL TWIN FOR POWER SYSTEMS

Creating a DT of power systems involves modelling and simulating physical assets, processes, or systems required for power generation, transmission, and distribution networks to provide a virtual replica capable of forecasting performance, enhancing operational efficiency, and assisting decision-making via real-time data and analytics [102]. Data federation is an approach that allows integration, unification, and governance of data stored in various sources by using a federated query engine that translates a single query into subqueries that are distributed to data sources for processing and analysis [103]. It can facilitate the unified access and analysis without requiring data to be duplicated or relocated [104], [105]. The primary benefits of data federation include instantaneous data access, minimized data storage and the ability to access data from various places, and eliminating movement of data. In the power systems DT context, data federation, a foundation to create efficient DT, is paramount for efficiently accessing and combining myriad data types from distributed sources, such as grid network information,

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3580055

operational and security logs, and external data like weather information. These sources are often heterogeneous in format and semantics. As illustrated in Figure 5, data federation offers a solution by providing a unified view of this disparate data, thus enabling DTs to access and analyse the information as if it were stored within a single, coherent database, eliminating the need for physical integration or data duplication.

While data federation enables seamless and unified access to data, it does not inherently address how to manage the functional complexity of a DT. This complexity arises from the diverse operational functionalities—fault detection, stability analysis, and predictive maintenance. Vargas et al. [44] proposed a modular framework emphasizing functional independence to address this. The operational capabilities of DTs are organized in this framework into distinct, task-specific modules that are independently developable, maintainable, and scalable. This modular approach depends on reliable and consistent data inputs, which the federated architecture can supply, even while it offers flexibility and efficiency for localized operations. For the modular design to be effective in applications like renewable integration, integrated, high-quality data must be available. By providing the essential inputs for each module to function well and acting as the basis for unified data access, federated architecture can be used in conjunction with the modular framework. For distribution throughout the larger network of power systems, the modular framework's outputs, like localized analysis or predictive insights, can be fed back into the federated system. However, the modular framework's scalability across entities is limited by its more focused emphasis on organizational adaptation. This emphasizes how the federated architecture serves as the fundamental layer that makes it possible for strong, cross-organizational DT systems to be assured in their expansion potential. Together, these approaches can produce a comprehensive, scalable, and effective DT architecture that takes into account both data interoperability and functional flexibility.

A. DATA TYPES

A successful operation of power systems DT relies on a comprehensive understanding and utilization of various data sources [15], [106]. On the other hand, data requirements for the digital twinning of power systems are extensive and varied based on the complexity and the specific objectives of the DT [84], [88]. Zhang et al. [15] categorize DT data into five primary data types, i.e., data related to physical entity and virtual entity, domain knowledge, data related to service, and fusion data. Furthermore, additional operational and analytical data of power systems DTs are also required [108]. The types of those data sources can also be varied. Firstly, structured data includes tabular formats like operational parameters, sensor readings, and configuration settings SCADA systems; and unstructured data comprises textual information from maintenance logs, incident reports, and documentation from vendors; time-series data produced by sensors, smart meters, and IoT devices. In addition,

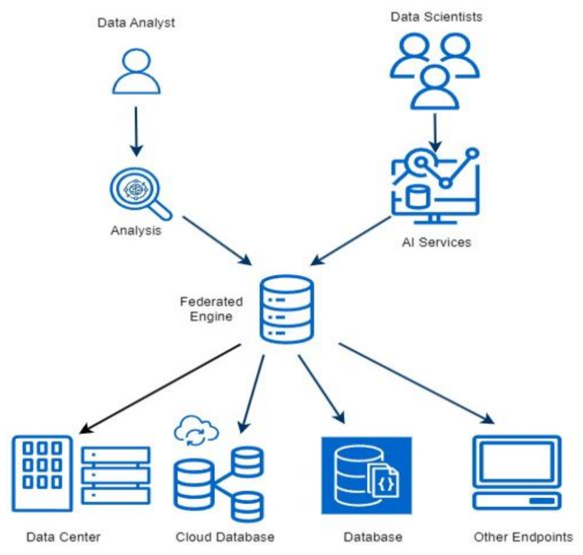

FIGURE 5. Overview of data federation.

geospatial data might detail the physical location and layout of infrastructure components such as power generation sources and transmission lines. Data can also be in graphs illustrating the relationships and dependencies between different constituent systems, such as the connectivity between nodes in the power grid. Ensuring the accuracy, timeliness, and security of this diverse data is pivotal for the successful implementation of DT and power system operations.

B. DATA ATTRIBUTES

Data attributes necessary for developing DTs of power systems can be categorized as in Figure 6.

1) CORE SYSTEM DATA

Operational Data i.e., time-synchronized grid measurements, sensor readings reflecting power output, voltage levels, frequency, rotor angle, current flows, and temperatures data required for real-time monitoring and decision-making.
Asset Performance, i.e., data on different equipment, assets and how they perform under various conditions, their efficiency, output, downtime, and more, which is vital for optimizing asset utilization and planning maintenance.
System Configuration and Network Data, i.e., grid models and topology data, geographic coordinates of the plant, plant type, detailed information on the power network's configuration, including grid connectivity and interdependencies.

2) ANALYTICAL AND PREDICTIVE DATA

Historical Performance and Trend Data, i.e., historical data, trend analysis, and equipment health records, which can be useful for benchmarking, predictive modelling, and maintenance optimization.
Simulation and Model Parameters, i.e., data used in simulations to test various scenarios and outcomes, enabling fine-tuning of the DT and predictive analysis of system behaviours under different conditions.

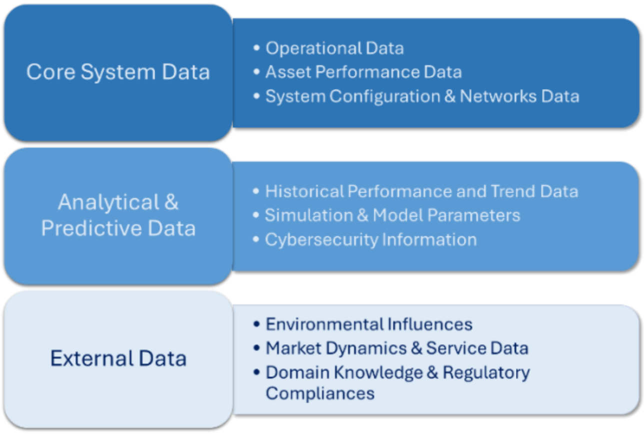

FIGURE 6. Different types of data attributes.

Cybersecurity Information, i.e., security logs and threat intelligence, including data on access, incident reports, vulnerability assessments, and information on potential cybersecurity threats. This data is vital for detecting, and analysing cyber threats and vulnerabilities, and ensuring the security and integrity of the power system and its DT.

3) EXTERNAL DATA

Environmental Data, i.e., weather and climate data (historical and forecasted) and geographical data and topological features, which are crucial for adapting system operations to environmental conditions.

Market Dynamics and Service Data, i.e., consumption patterns, demand forecasts, compliance with regulations and standards, and market data like energy prices.

Domain Knowledge and Regulatory Compliance, i.e., best practices, maintenance strategies, expert experience, predefined rules, industry standards, industry guidelines.

C. DATA REQUIREMENTS AND PRINCIPLES

In DT, addressing various data requirements, guided by data principles, is pivotal for enhancing DTs' accuracy, efficiency, and adaptability. Zhang et al. [15] discussed different data requirements and the principles for DT data.

Comprehensive data collection is essential, aiming to capture a full spectrum of conditions and events, supported by the *"complementary principle,"* which advocates for integrating physical and virtual data sources. This is complemented by the necessity for real-time interaction, where immediate data exchanges between physical and virtual models are facilitated for dynamic adjustments, guided by the *"timeliness principle"* to ensure rapid data synchronization. The universality of data across different DT scenarios is achieved through the *"standardized principle,"* which promotes the adoption of uniform data formats to facilitate universal application. Furthermore, knowledge mining, which extracts valuable insights from data to refine virtual models, is underpinned by the *"association principle",* which focuses on identifying data relationships. Data fusion is another aspect, combing data from various sources to enhance overall data quality, with the *"fusion principle"* emphasizing the merging of diverse data sets for a more comprehensive analysis. Iterative optimization plays a significant role in continuously improving data quality through repeated fusion and analysis, driven by the *"information growth principle,"*, which evaluates and enhances the information content of data. Lastly, accessible usage aims to simplify data access for users of varying expertise levels. This is achieved through the *"servitization principle,"* which involves packaging data and data-related resources into on-demand services.

These principles and approaches collectively underline the importance of a approach to data management in DTs, ensuring that the technology can be applied across a wide range of scenarios with enhanced precision and adaptability. Notably, the fusion and standardized principles underscore the need for integrating diverse data sources, maintaining data consistency, and promoting seamless data sharing across different platforms and applications, highlighting the importance of data integration approaches like data federation [84].

D. COMPONENTS FOR DATA FEDERATION

1) CANONICAL DATA MODELS (CDMS)

CDMs streamline the process of integrating various systems and databases by standardizing data entities and relationships into a simplified, universal format. This approach aims to establish a common language for managing data across different systems, which typically operate with their own unique languages, syntaxes, and protocols. The essence of a CDM lies in its ability to provide a unified definition of data, facilitating easier integration between systems, leading to improved operational processes, practices, and simplified data analytics. CDM introduces a new, distinct model that can encapsulate and translate diverse types of data [109].

2) KNOWLEDGE GRAPH

Contextualization is the process of creating meaningful linkages between data sources and types so users may navigate and find data. Through contextualization, knowledge graphs can be built [110]. Knowledge graph represents various entities, such as components of the electrical grid, and delineates the relationships between them, effectively mapping out the intricate web of connections that constitute the grid. Utilizing domain-specific ontology based on electrical infrastructure, the knowledge graph incorporates diverse properties of each entity, offering a detailed understanding of the grid's components and their interrelations and allows for the representation of complex concepts and relationships within the electrical infrastructure [2]. By capturing and organizing data in this manner, the knowledge graph serves as an invaluable resource for various stakeholders and engineers to help them to deeply understand the target system, facilitating the establishment of a common knowledge base [111].

3) VERIFICATION MODULE

Maintaining high-quality, accurate, and secure data is paramount when aggregating and providing unified data

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3580055

access. The data verification engine ensures the data model's quality is assessed and maintained throughout the integration process and should be designed to implement rigorous quality checks and validation processes. That includes identifying and correcting any inaccuracies, or anomalies in the aggregated data, ensuring adherence to predefined standards and schemas, and error correction routines to maintain federated data's the reliability and integrity [2]. To sustain data quality over time, it also should implement robust data governance practices, establish protocols for data management and continuous quality control, and enforce stringent security measures, including encryption, and access controls, complemented by regular security audits to protect unauthorized access and ensure integrity and trustworthiness within a secure federation environment.

The approaches to ensure the effective operation of DTs, and the validity and reliability of data at both the individual and federated levels are proposed in [111]. Data validation process for the DT includes analyzing the characteristics and patterns of input data, such as time-series data, sporadic event data, or anomaly-sensing data. This analysis can identify efficacy criteria, such as thresholds for "normal," "caution," and "warning" zones, and identifies data features. These criteria can be established using statistical techniques, but their accuracy and application are improved by combining advanced causality analysis techniques, like Recurrent Neural Networks (RNNs), with expert domain knowledge. The authors emphasized the important of real-time validation to ensure the time-series data stays within acceptable ranges. Moreover, by integrating various data types, such as weather and temperature data, among others, the authors also suggested the use of attribute-based validation tools for assessing the system's effectiveness.

4) CUSTOMIZED ADAPTERS OR STANDARDIZED INTERFACES

In order to bring the disparate data silos together, they are combined into a SSoT or data federation system using either a specially designed adapter or standardized interface [2]. The implementation of standardized interfaces or the development of customized adapters enables access to a broad range of heterogeneous data models. It also facilitates seamless communication and data exchange across a diverse array of systems by harmonizing data formats, protocols, and communication [113]. For instance, FIWARE provides a set of Application Programming Interfaces (APIs) that can be used for integration of IoT components that adopt communication protocol the FIWARE platform supports [114]. Through FIWARE Next Generation Service Interfaces (NGSI) API, system components can interface with Orion Context Broker, which maintains updated context information from components and applications, after receiving data from devices and gateways. Similarly, SQL-based APIs, for instance, Java Database Connectivity (JDBC) [115], and Open Database Connectivity (ODBC) [116] are designed to access and adapt heterogeneous data sources to a relational model, often by transforming complex data structures into a flattened format. These adapters ensure that data from varied sources is accurately represented, up-to-date, and accessible within the SSoT, thereby enhancing data consistency, reducing integration complexity.

VI. DATA INTEROPERABILITY FOR FEDERATION AND CHALLENGES

A. DATA MODELS

For the federation of data within power systems, seamless interoperability measures are essential for communicating and exchanging data among a range of devices, systems, and stakeholders. Data interoperability provides the capabilities required for data exchange, including data models, data formats, and interfaces. In this regard, many organizations have developed standards and frameworks to ensure data interoperability, enabling various constituent systems in the energy domain to communicate and interoperate effectively [117]. For example, SAREF (Smart Applications REFerence Ontology), Gaia-X, SGAM, and CIM have been adopted to provide guidelines to avoid data silos and integrate and interpret data for seamless collaboration and communication in energy data space. Data modeling is a basis for data federation and enables more efficient data architecture planning. In the context of DTs, which encompasses metadata, condition or state, event data, and analytics, data modeling becomes a key component in converting siloed data into scalable solutions. Data modeling in software engineering refers to the process of simplifying the representation of a software system's diagram or data model through the application of specific formal methods, which involves describing data using a combination of textual descriptions and symbolic representations. Data modelling, according to IBM, is the process of visualizing an information system to convey relationships between data elements and structures [118]. Data modeling can be achieved through visual representations that detail attributes, relationships, and data storage locations. There are distinct types of data models - conceptual, logical, and physical data models, and data modeling techniques like Entity-Relationship (ER) diagrams [119], UML class diagrams [120], and data dictionaries [121] have been utilized for abstracting the relationships between different data entities and visualizing how components of a physical entity and its operational data are interconnected.

B. OPEN STANDARDS FOR DATA INTEROPERABILITY BETWEEN HETERONEOUS POWER SYSTEMS

1) SAREF

Published as a set of open standards by the European Telecommunications Standards Institute (ETSI) Technical Committee Smart Machine to Machine Communications, Applications REFerence Ontology (SAREF) provides a suite of ontologies that forms a shared model for semantic interoperability between different sectors in the IoTs and

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3580055

contributes to data space development. SAREF comprises a core ontology and its extended ontologies for different domains, and two of them that can contribute to the energy domain are SAREF4ENER and SAREF4GRID [122]. A standardized ontology, SAREF4ENER, is used to represent data in the energy domain and facilitate communication between energy-related information systems. This includes connecting disparate data models. SAREF4GRID is used to for the domain of the electrical grid to enable data sharing and interoperability among various grid-related systems and devices.

To demonstrate the practical implementation of the SAREF ontology and validate the ontology, Weerdt et al. [122] validated the ontology by expressing all the information available from different smart devices in a home and enabling interoperability by allowing communication between smart devices. In their study, the authors mapped IoT data from smart devices into a knowledge graph format and tested the capability to effectively represent data from real devices. In addition, they developed an IoT setup using Raspberry Pi devices to simulate interactions between a thermometer, thermostat, and heater. All communications used SAREF to demonstrate interoperability. The study showed that SAREF supports modular extensions for domain-specific needs and demonstrated the capacity to integrate different devices under a unified framework. Its ability to connect diverse devices can also be applied to similar projects requiring standardized communication in IoT environment.

2) SGAM

The working group of EU Mandate M/490's Reference Architecture produced the Smart Grid Architecture Model (SGAM) to offer an approach for developing Smart Grid architectures. SGAM is a framework developed to tackle modern energy systems' growing complexity and interoperability challenges. It covers the interoperability between systems or components of the energy chain, from generation, transmission, distribution, and customer premises [123]. SGAM comprises five interoperability levels representing business processes, functions, communication protocols, objectives, information exchange and models, and components. It also includes diverse aspects, such as the information flows between technical functions, the components that carry out the technical functions in the system, and the standard protocols and data models that facilitate these information flows. An SGAM-based approach to analyze smart grid solutions in the DISCERN European research project has been proposed [124]. The DISCERN project evaluated smart grid solutions across different DSO implementations, using SGAM and IEC 62559 to support knowledge sharing and solution adaptation. It demonstrated enhanced monitoring of Medium Voltage/Low Voltage Networks by mapping use cases into SGAM layers, providing consistent representation and comparison of solutions. It also proposed a web-based tool that supports the collaborative development of use cases and SGAM models while enabling automatic analysis, such as requirements extraction, 3D visualization of architectures, and component identification. The DISCERN project demonstrated the potential of SGAM and IEC 62559 in fostering interoperable and standardized smart grid solutions. Developing the SGAM use cases to deal with other smart grid challenges, such as Electric Vehicles (EV) integration and microgrid management, can be potential applications of SGAM. In addition, future implementations can also consider broadening the SGAM's applicability beyond pilot projects to national or regional grid systems and facilitating dynamic updates of SGAM libraries based on real-world implementations.

3) GAIA-X

Gaia-X is another European initiative that creates an interoperable, decentralized, data federated, and secure infrastructure [126]. Gaia-X is used to specify the requirements, design the architecture, and implement the software components to connect multiple stakeholders in a federation. It ensures interoperability, transparency, data security, and controllability of services through a standard description format, identity management, and compliance verification mechanisms. While enabling flexibility to adapt to industry-specific requirements, Gaia-X supports interoperable open interfaces. It also enables secure data exchange and decentralized storage where all data remains in the storage of organizations, thereby data owners have complete control of their data [127].

Gaia-X-based data spaces consider regulations and policy guidelines regarding the gathering, storing, and using of data, such as the EU Data Act. Through its federated data spaces and ecosystems, it plays an important role in supporting compliance with these regulations, providing a solid foundation for data management. Forte et al. [127] demonstrated how Gaia-X has been used in the industry. The manufacturing process of GMN (German Mechanical Engineering Company) used Gaia-X concepts. During quality testing, it uses cutting-edge sensor technology to generate digital fingerprints for motor spindles. To facilitate secure data sharing with customers, it also made these fingerprints available as datasets in ecosystems that comply with Gaia-X. It also makes new data-driven services like extended maintenance options, end-of-life forecasts, and remote diagnostics possible. To improve communication between IIoT platforms, Manufacturing Execution Systems (MES), and Product Lifecycle Management (PLM), Gaia-X principles were also incorporated into a data management platform. It facilitates seamless aggregation and publishing of data and improves operational capabilities and collaboration. The Gaia-X facilitates scalable data ecosystems by standardizing governance and information exchange formats. The aforementioned use case demonstrates how manufacturers like GMN have added direct value to consumers by using Gaia-X ecosystems to provide customized, data-driven services like spindle health monitoring. While complying with the regulatory requirements, it is still difficult to ensure compliance with various regional regulations, particularly in global deployments, and to modify current IT infrastructures to meet Gaia-X requirements, such as securing data handling.

4) CIM

Using industrial standards such as the canonical -based Grid Model Exchange Standard profile facilitates the sharing of updates to current and future grid models among external stakeholders, promoting vendor interoperability within the energy sector [129]. This International Electrotechnical Commission (IEC)'s CIM standard comprising a Specification, Schema, and Metamodel, developed by the Distributed Management Task Force (DMTF), is integral to the Web-Based Enterprise Management (WBEM) initiative. This initiative founds a unified framework for managing information across systems, networks, applications, and services [130]. For instance, the European Network of Transmission System Operators for Electricity (ENTSO-E) highlights that CIM for grid model exchange can enable the exchange of data vital for local or European-wide grid development research. The process of exchanging grid models encapsulates a wide array of applications, such as sharing information about power system equipment, grid topology, state variables of the power system, steady-state assumptions, and facilitating the management and analysis of market data, contingency analysis, and dynamic security assessments [131]. CIM employs foundational technologies such as the Unified Modeling Language (UML), the eXtensible Markup Language (XML), and the Resource Description Framework (RDF) to model, exchange, and ensure interoperability of data [58].

The most essential standards to this framework are the IEC 61970 and IEC 61968 standards, which are integral for the exchange of information in both transmission and distribution grids [132]. CIM includes classes that have specific attributes to represent different data object types needed for data exchange between Transmission System Operators (TSOs) and Distribution System Operators (DSOs) [133]. It is a common language that abstracts the specifics of each system's data model into a unified framework. This standardization allows for directly mapping different data elements and structures into a format understood by all parties despite the original differences in data representation. CIM's adaptability is a key feature, allowing for extensions to cover specific demands and the flexibility to design data exchange profiles constructed from CIM and custom extensions as a subset of semantic canonical model. Manufacturers, TSOs, and DSOs can construct their own CIM profiles, including all or part of the standardized CIM model, to meet their needs for data modelling [134]. For instance, to establish a common language to interoperate and common messaging between systems, CIM has been adopted as the reference data model in different research projects such as EU SysFlex [135], TDX-ASSIST [136], and OneNet [137] for various application cases. Those use cases are related to data exchanges and interfaces including transferring, anonymizing and aggregating energy data, prediction of production and consumption for operation planning, management of active power flexibility for congestion and voltage control.

To enhance smart grid automation, Naumann et al. [137] proposed the integration of two important standards, IEC 61850 and IEC 61970/61968 (CIM), with an emphasis on protective systems. Both standards focus on distinct aspects of the smart grid to facilitate automatic and flexible protection methods. The author suggested a methodology for mapping IEC 61850 data formats to CIM objects to facilitate integration across various grid management levels. For instance, CIM classes for analog measurements and protection characteristics were mapped to logical nodes (LNs) in IEC 61850, such as MMXU (measurement) and PTOC (protection relay). Custom extensions are necessary as CIM does not have predefined models for some protective functions. In this regard, the authors expanded the CIM and created unique CIM classes to model protective functionality that is not yet standardized. For example, new CIM objects were mapped to characteristics of overcurrent protection. The study demonstrates CIM's adaptability to user-specific adaptations and capacity to satisfy changing grid needs. However, scaling the integrated framework to larger grids with various device types can still be challenging. The development of automatic ontology mapping tools between IEC 61850 and CIM may help to reduce errors and manual efforts. Future research can focus on incorporating features like real-time data synchronization between standards and semantic evaluation. Another research direction could be to test the integrated framework in situations with high renewable penetration, EV charging networks, or microgrids to assess performance under various circumstances.

C. DATA FEDERATION CHALLENGES

1) CONNECTIVITY AND INTEGRATION

The physical network of the electric grid, with its diverse components, is an important part to enable the smooth data from the grid's physical layer to its DT, using various communication protocols. However, this data exchange comes with its challenges. Synchronization issues can occur when devices fail to receive the necessary data or signal network disruptions. Despite the advancements in IoT technologies and the deployment of 5G networks, which enhance connectivity within the DT framework, complexities such as software errors, updates, and latency issues persist, potentially hampering real-time monitoring and accuracy of data [6].

Furthermore, as the significance of sophisticated data analytics and management systems grows, power system DTs must not only handle the voluminous data generated from diverse sources but also ensure its integrity, accuracy,

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2025.3580055

and timeliness [7]. The integration and federation of disparate data types, often in incompatible formats, pose a significant barrier to achieving a seamless and secure flow of information [8]. However, overcoming these challenges could lead to significant benefits, such as optimizing grid performance and enabling advanced analytics to predict demand, manage supply, and mitigate potential disruptions in real time. The urgency of addressing these connectivity and data integration challenges cannot be overstated. Edge federated ML approaches have been popular for training the ML model by using data gathered locally on edge devices, and updating the global model in a central server, thereby reducing latency, resource utilization and improving bandwidth availability [139].

2) STANDARDIZATION

Deployment of DT necessitates a uniform framework to define, store, and execute DT models that can ensure interoperability and seamless integration across different systems. As energy systems contain diverse and heterogeneous system components from generation and transmission to distribution and consumption, they require a unified cross-system collaboration and interaction platform.

While standardization efforts and protocols have been in place to ensure interoperability for data federation in DT, they have limitations. CIM standard, for instance, offers a comprehensive and detailed schema covering different energy utility appliances. However, extensive customization is required to fit the CIM into specific operational scenarios. That can lead to longer deployment times and result in deviations in how different organizations implement CIM, potentially affecting interoperability. Similarly, the Generic Object-Oriented Substation Event (GOOSE) protocol, standardized as IEC 61850, has also been widely utilized for real-time data exchange in digital substations. Despite its numerous benefits, the GOOSE has limitations like security vulnerabilities since it did not extensively focus on cybersecurity measures [140]. That could lead to GOOSE communication and messages susceptible to different types of cyber-attacks when proper security control and protection mechanisms are not deployed. Moreover, since the energy sector is heavily regulated, compliance with local, national, and international regulations and standardizing DT deployment across different regulatory environments could also be an issue because local regulations may dictate specific requirements for data handling, system safety, and operational procedures.

3) DATA MANAGEMENT AND GOVERNANCE

The success of a DT relies on its underlying data quality. Consistent and high-quality data streams are not just important; they are crucial for DTs to function optimally. Poor and inconsistent data can significantly impair a DT's functionality and its ability to optimize power system operations [14], [15], [56]. This underscores the need for meticulous planning, data generation, collection, and management efforts to ensure the capture of relevant and high-quality data across power generation, transmission, and distribution process.

The complexities of data ownership and governance pose significant challenges in deploying DTs [57], prompting questions about the rights to share specific data. Even though a DT ecosystem development necessitates collaboration and data exchange between system stakeholders, they handle critical infrastructure, and the data they manage includes sensitive and confidential information that could potentially expose the grid to severe security risks if disclosed. For potential collaborative research and development aiming for the development of new methodologies, limited access to real-world operational data is a significant barrier. Thus, it could be formidable to test hypotheses accurately, validate models, or simulate realistic scenarios that would provide meaningful results. Despite the use of synthetical data generation techniques to overcome the limitations, how to carefully generate adequate and high-quality data to accurately reflect the complex dynamics of actual system operations and be appropriate for training ML models remains a prerequisite [141].

4) DATA SECURITY

Data security can be another challenge to the operational integrity of DT systems. As DTs provide a digital-physical association, the security of the data link connecting them becomes imperative since it exchanges critical data between the virtual and physical world, which inherently possesses vulnerabilities and risks to data breaches, corruption, unauthorized access, and cyber threats [142]. Consequently, ensuring the security of communication medium is paramount, requiring rigorous compliance with data security requirements such as privacy, authentication, integrity, and traceability throughout the development of DTs. Implementing vigorous data security measures, including encryption of data, controlled access privileges, penetration testing, and source code scanning becomes essential in mitigating potential vulnerabilities [6], [70], [72]. Furthermore, emerging technologies such as blockchain can be helpful for DT communication security, ensuring data privacy and fostering trust within DT ecosystems [80]. These advancements signify a transformation towards more secure and resilient DT frameworks capable of resisting the evolving landscape of cyber threats.

VII. CONCLUSIONS

A significant proliferation in DERs and IoT devices drives the need for more streamlined data integration and exchange between utilities and energy systems. A DT of the power system serves as a virtual representation of physical infrastructure, enabling real-time reflection of system behavior through bidirectional data flow. DTs can enable interoperable and secure data exchange necessary for managing modern power systems by leveraging data federation principles. This review provided a power system-centric synthesis of DT developments, distinguishing it from broader reviews that often cover general energy systems or manufacturing applications. It identified the concept of DTs, current practices of DTs in power system context, the functional and non-functional requirements and how DT can satisfy its intended purpose and provide the expected outcomes. Unlike existing studies, this analysis investigated the concept and various aspects of data federation. It highlights requirements for data federation, data types, attributes, and principles for effective twinning. It then explores the supporting digital technologies for the digital twinning of power systems and how they are integrated into DT implementation by providing case studies. Another key contribution is the discussion of industrial interoperability standards and challenges, which are often underrepresented in existing literature. It presents an analysis of interoperability standards and highlights insights into their applicability and gaps. Through real-world examples and case studies, this review offers a practical lens that brings the operationalization of DTs into the power system context.

The insights provided in this study can provide a foundation for researchers, especially DT application developers in power system engineering, to delve deeper into this field. Future work will focus on defining domain-specific DT use cases, identifying required functionalities, and detailing key information modeling elements, particularly based on CIM. Through continued exploration and technological advancement, the full potential of DTs to transform power system operation and resilience can be realized, further driving the capabilities of smart grid technologies and sustainable energy solutions. Moreover, insights from this domain may serve as a foundation for extending DT applications into related areas such as Positive Energy Districts (PEDs), Positive Energy Buildings (PEBs), and community energy systems, where integrated energy management and interoperability are equally important.

ACKNOWLEDGMENT

This work has received funding from the European Commission in Horizon Europe TwinEU project under Grant Agreement 101136119.

REFERENCES

[1] "Digital twins and their use in future power systems | Digital Twin." Accessed: Feb. 04, 2024. [Online]. Available: https://digitaltwin1.org/articles/1-4

[2] C. Brosinsky, M. Naglič, S. Lehnhoff, R. Krebs, and D. Westermann, "A Fortunate Decision That You Can Trust: Digital Twins as Enablers for the Next Generation of Energy Management Systems and Sophisticated Operator Assistance Systems," *IEEE Power and Energy Mag.*, vol. 22, no. 1, Art. no. 1, Jan. 2024, doi: 10.1109/MPE.2023.3330120.

[3] "The Development of Digital Twin Technology Review | IEEE Conference Publication | IEEE Xplore." Accessed: May 02, 2024. [Online]. Available: https://ieeexplore-ieee-org.tudelft.idm.oclc.org/document/9327756

[4] "Digital Twin Techniques for Power Electronics-Based Energy Conversion Systems: A Survey of Concepts, Application Scenarios, Future Challenges, and Trends | IEEE Journals & Magazine | IEEE Xplore." Accessed: May 02, 2024. [Online]. Available: https://ieeexplore-ieee-org.tudelft.idm.oclc.org/document/9950705

[5] A. Kummerow, C. Monsalve, D. Rosch, K. Schafer, and S. Nicolai, "Cyber-physical data stream assessment incorporating Digital Twins in future power systems," in *2020 International Conference on Smart Energy Systems and Technologies (SEST)*, Istanbul, Turkey: IEEE, Sep. 2020, pp. 1–6. doi: 10.1109/SEST48500.2020.9203270.

[6] M. Atalay and P. Angin, "A Digital Twins Approach to Smart Grid Security Testing and Standardization," Jun. 2020, pp. 435–440. doi: 10.1109/MetroInd4.0IoT48571.2020.9138264.

[7] T. Wagner, C. Kittl, J. Jakob, J. Hiry, and U. Häger, "Digital Twins in Power Systems: A Proposal for a Definition," *IEEE Power and Energy Magazine*, vol. 22, no. 1, pp. 16–23, Jan. 2024, doi: 10.1109/MPE.2023.3328581.

[8] Md. M. H. Sifat *et al.*, "Towards electric digital twin grid: Technology and framework review," *Energy and AI*, vol. 11, p. 100213, Jan. 2023, doi: 10.1016/j.egyai.2022.100213.

[9] C. Vergara, R. Bahsoon, G. Theodoropoulos, W. Yanez, and N. Tziritas, "Federated Digital Twin," in *2023 IEEE/ACM 27th International Symposium on Distributed Simulation and Real Time Applications (DS-RT)*, Singapore, Singapore: IEEE, Oct. 2023, pp. 115–116. doi: 10.1109/DS-RT58998.2023.00027.

[10] M. Jafari, A. Kavousi-Fard, T. Chen, and M. Karimi, "A Review on Digital Twin Technology in Smart Grid, Transportation System and Smart City: Challenges and Future," *IEEE Access*, vol. 11, pp. 17471–17484, 2023, doi: 10.1109/ACCESS.2023.3241588.

[11] C. Alcaraz and J. Lopez, "Digital Twin: A Comprehensive Survey of Security Threats," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 3, pp. 1475–1503, 2022, doi: 10.1109/COMST.2022.3171465.

[12] D.-Y. Jeong *et al.*, "Digital Twin: Technology Evolution Stages and Implementation Layers With Technology Elements," *IEEE Access*, vol. 10, pp. 52609–52620, 2022, doi: 10.1109/ACCESS.2022.3174220.

[13] S. Mihai *et al.*, "Digital Twins: A Survey on Enabling Technologies, Challenges, Trends and Future Prospects," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 4, pp. 2255–2291, 2022, doi: 10.1109/COMST.2022.3208773.

[14] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital Twin: Enabling Technologies, Challenges and Open Research," *IEEE Access*, vol. 8, pp. 108952–108971, 2020, doi: 10.1109/ACCESS.2020.2998358.

[15] "Digital twin data: methods and key technologies | Digital Twin." Accessed: Mar. 13, 2024. [Online]. Available: https://digitaltwin1.org/articles/1-2

[16] M. Grieves, *Origins of the Digital Twin Concept*. 2016. doi: 10.13140/RG.2.2.26367.61609.

[17] M. Eckhart and A. Ekelhart, "Digital Twins for Cyber-Physical Systems Security: State of the Art and Outlook," in *Security and Quality in Cyber-Physical Systems Engineering*, S. Biffl, M. Eckhart, A. Lüder, and E. Weippl, Eds., Cham: Springer International Publishing, 2019, pp. 383–412. doi: 10.1007/978-3-030-25312-7_14.

[18] M. Grieves and J. Vickers, "Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems," in *Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches*, F.-J. Kahlen, S. Flumerfelt, and A. Alves, Eds., Cham: Springer International Publishing, 2017, pp. 85–113. doi: 10.1007/978-3-319-38756-7_4.

[19] Z. Shen, F. Arraño-Vargas, and G. Konstantinou, "Virtual testbed for development and evaluation of power system digital twins and their applications," *Sustainable Energy, Grids and Networks*, vol. 38, p. 101331, Jun. 2024, doi: 10.1016/j.segan.2024.101331.

[20] D.-G. J. Opoku, S. Perera, R. Osei-Kyei, and M. Rashidi, "Digital twin application in the construction industry: A literature review," *Journal of Building Engineering*, vol. 40, p. 102726, Aug. 2021, doi: 10.1016/j.jobe.2021.102726.

[21] H. Chen, Z. Zhang, P. Karamanakos, and J. Rodriguez, "Digital Twin Techniques for Power Electronics-Based Energy Conversion Systems: A Survey of Concepts, Application Scenarios, Future Challenges, and Trends," *IEEE Industrial Electronics Magazine*, vol. 17, no. 2, pp. 20–36, Jun. 2023, doi: 10.1109/MIE.2022.3216719.

[22] M. Lamagna, D. Groppi, M. M. Nezhad, and G. Piras, "A comprehensive review on digital twins for smart energy management system," *Int. J. EQ*, vol. 6, no. 4, pp. 323–334, Nov. 2021, doi: 10.2495/EQ-V6-N4-323-334.

[23] A. G. Abo-Khalil, "Digital twin real-time hybrid simulation platform for power system stability," *Case Studies in Thermal Engineering*, vol. 49, p. 103237, Sep. 2023, doi: 10.1016/j.csite.2023.103237.

[24] W. Danilczyk, Y. Sun, and H. He, "ANGEL: An Intelligent Digital Twin Framework for Microgrid Security," in *2019 North American*

Power Symposium (NAPS), Oct. 2019, pp. 1–6. doi: 10.1109/NAPS46351.2019.9000371.

[25] J. Huang, L. Zhao, F. Wei, and B. Cao, "The Application of Digital Twin on Power Industry," *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 647, no. 1, p. 012015, Jan. 2021, doi: 10.1088/1755-1315/647/1/012015.

[26] N. Tzanis, N. Andriopoulos, A. Magklaras, E. Mylonas, M. Birbas, and A. Birbas, "A Hybrid Cyber Physical Digital Twin Approach for Smart Grid Fault Prediction," in *2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS)*, Jun. 2020, pp. 393–397. doi: 10.1109/ICPS48405.2020.9274723.

[27] J. Bofill, M. Abisado, J. Villaverde, and G. A. Sampedro, "Exploring Digital Twin-Based Fault Monitoring: Challenges and Opportunities," *Sensors*, vol. 23, no. 16, Art. no. 16, Jan. 2023, doi: 10.3390/s23167087.

[28] P. Jain, J. Poon, J. P. Singh, C. Spanos, S. R. Sanders, and S. K. Panda, "A Digital Twin Approach for Fault Diagnosis in Distributed Photovoltaic Systems," *IEEE Trans. Power Electron.*, vol. 35, no. 1, pp. 940–956, Jan. 2020, doi: 10.1109/TPEL.2019.2911594.

[29] "Fault prediction in transmission line using GOOSE protocol and Artificial Neural Network | IEEE Conference Publication | IEEE Xplore." Accessed: May 02, 2024. [Online]. Available: https://ieeexplore-ieee-org.tudelft.idm.oclc.org/document/9616801

[30] Y. Peng, S. Zhao, and H. Wang, "A Digital Twin Based Estimation Method for Health Indicators of DC–DC Converters," *IEEE Transactions on Power Electronics*, vol. 36, no. 2, pp. 2105–2118, Feb. 2021, doi: 10.1109/TPEL.2020.3009600.

[31] P. Jia, X. Wang, and X. Shen, "Digital-Twin-Enabled Intelligent Distributed Clock Synchronization in Industrial IoT Systems," *IEEE Internet Things J.*, vol. 8, no. 6, pp. 4548–4559, Mar. 2021, doi: 10.1109/JIOT.2020.3029131.

[32] A. Barni, A. Fontana, S. Menato, M. Sorlini, and L. Canetta, "Exploiting the Digital Twin in the Assessment and Optimization of Sustainability Performances," in *2018 International Conference on Intelligent Systems (IS)*, Sep. 2018, pp. 706–713. doi: 10.1109/IS.2018.8710554.

[33] "Cemig Group," Cemig. Accessed: May 02, 2024. [Online]. Available: https://www.cemig.com.br/en/grupo-cemig

[34] "Digital twins at CERN and beyond," CERN. Accessed: May 02, 2024. [Online]. Available: https://home.cern/news/news/knowledge-sharing/digital-twins-cern-and-beyond

[35] "ABB Group. Leading digital technologies for industry," ABB Group. Accessed: May 02, 2024. [Online]. Available: https://global.abb/group/en

[36] "CORYS | DYNAMIC SIMULATION," CORYS. Accessed: May 02, 2024. [Online]. Available: https://www.corys.com/en/

[37] "Electrical Digital Twin Brochure".

[38] "Welcome to State Grid Corporation of China." Accessed: May 02, 2024. [Online]. Available: http://www.sgcc.com.cn/html/sgcc_en/index.shtml

[39] "gPROMS Digital Process Design and Operation," siemens.com Global Website. Accessed: May 02, 2024. [Online]. Available: https://www.siemens.com/global/en/products/automation/industry-software/gproms-digital-process-design-and-operations.html

[40] "Digital Twin | Siemens Software," Siemens Digital Industries Software. Accessed: May 02, 2024. [Online]. Available: https://www.sw.siemens.com/en-US/technology/digital-twin/

[41] B. Gandhi, "Will generative AI make the digital twin promise real in the energy and utilities industry?," IBM Blog. Accessed: May 02, 2024. [Online]. Available: https://www.ibm.com/blog/will-generative-ai-make-the-digital-twin-promise-real-in-the-energy-and-utilities-industry/www.ibm.com/blog/will-generative-ai-make-the-digital-twin-promise-real-in-the-energy-and-utilities-industry

[42] "(주)테크트리이노베이션." Accessed: May 02, 2024. [Online]. Available: https://tt-inno.com/

[43] *Non-Functional Requirements in Software Engineering*. Accessed: May 02, 2024. [Online]. Available: https://link.springer.com/book/10.1007/978-1-4615-5269-7

[44] F. Arraño-Vargas and G. Konstantinou, "Modular Design and Real-Time Simulators Toward Power System Digital Twins Implementation," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 52–61, Jan. 2023, doi: 10.1109/TII.2022.3178713.

[45] M. Glatt, C. Sinnwell, L. Yi, S. Donohoe, B. Ravani, and J. C. Aurich, "Modeling and implementation of a digital twin of material flows based on physics simulation," *Journal of Manufacturing Systems*, vol. 58, pp. 231–245, Jan. 2021, doi: 10.1016/j.jmsy.2020.04.015.

[46] F. Tao, B. Xiao, Q. Qi, J. Cheng, and P. Ji, "Digital twin modeling," *Journal of Manufacturing Systems*, vol. 64, pp. 372–389, Jul. 2022, doi: 10.1016/j.jmsy.2022.06.015.

[47] A. Ait-Alla, M. Kreutz, D. Rippel, M. Lütjen, and M. Freitag, "Simulation-based Analysis of the Interaction of a Physical and a Digital Twin in a Cyber-Physical Production System," *IFAC-PapersOnLine*, vol. 52, no. 13, pp. 1331–1336, Jan. 2019, doi: 10.1016/j.ifacol.2019.11.383.

[48] S. Boschert and R. Rosen, "Digital Twin—The Simulation Aspect," in *Mechatronic Futures: Challenges and Solutions for Mechatronic Systems and their Designers*, P. Hehenberger and D. Bradley, Eds., Cham: Springer International Publishing, 2016, pp. 59–74. doi: 10.1007/978-3-319-32156-1_5.

[49] Y. Cui, F. Xiao, W. Wang, X. He, C. Zhang, and Y. Zhang, "Digital Twin for Power System Steady-state Modelling, Simulation, and Analysis," in *2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2)*, Oct. 2020, pp. 1233–1238. doi: 10.1109/EI250167.2020.9346850.

[50] F. Naseri *et al.*, "Digital twin of electric vehicle battery systems: Comprehensive review of the use cases, requirements, and platforms," *Renewable and Sustainable Energy Reviews*, vol. 179, p. 113280, Jun. 2023, doi: 10.1016/j.rser.2023.113280.

[51] "Toward Intelligent Cyber-Physical Systems: Digital Twin Meets Artificial Intelligence | IEEE Journals & Magazine | IEEE Xplore." Accessed: May 02, 2024. [Online]. Available: https://ieeexplore-ieee-org.tudelft.idm.oclc.org/document/9530501

[52] Q. Qiao, J. Wang, L. Ye, and R. X. Gao, "Digital Twin for Machining Tool Condition Prediction," *Procedia CIRP*, vol. 81, pp. 1388–1393, Jan. 2019, doi: 10.1016/j.procir.2019.04.049.

[53] Z. Shen, F. Arraño-Vargas, and G. Konstantinou, "Artificial intelligence and digital twins in power systems: Trends, synergies and opportunities," *Digital Twin*, vol. 2, p. 11, Aug. 2022, doi: 10.12688/digitaltwin.17632.1.

[54] B. Xu *et al.*, "A case study of digital-twin-modelling analysis on power-plant-performance optimizations," *Clean Energy*, vol. 3, no. 3, pp. 227–234, Nov. 2019, doi: 10.1093/ce/zkz025.

[55] "Cyber–Physical Security of a Smart Grid Infrastructure | IEEE Journals & Magazine | IEEE Xplore." Accessed: Aug. 13, 2024. [Online]. Available: https://ieeexplore-ieee-org.tudelft.idm.oclc.org/abstract/document/6016202

[56] M. Manngård, W. Lund, and J. Björkqvist, "Using Digital Twin Technology to Ensure Data Quality in Transport Systems," Mar. 2020.

[57] "Data management for developing digital twin ontology model." Accessed: Mar. 05, 2024. [Online]. Available: https://journals.sagepub.com/doi/epub/10.1177/0954405420978117

[58] M. Uslar, M. Specht, S. Rohjans, J. Trefke, and J. M. González, *The Common Information Model CIM: IEC 61968/61970 and 62325 - A practical introduction to the CIM*. Springer Science & Business Media, 2012.

[59] B. A. Talkhestani *et al.*, "An architecture of an Intelligent Digital Twin in a Cyber-Physical Production System," *at - Automatisierungstechnik*, vol. 67, no. 9, pp. 762–782, Sep. 2019, doi: 10.1515/auto-2019-0039.

[60] "REAL-TIME SIMULATION IN MANUFACTURING SYSTEMS: CHALLENGES AND RESEARCH DIRECTIONS | IEEE Conference Publication | IEEE Xplore." Accessed: Aug. 15, 2024. [Online]. Available: https://ieeexplore-ieee-org.tudelft.idm.oclc.org/document/8632542

[61] "Digital Twin Platforms: Requirements, Capabilities, and Future Prospects | IEEE Journals & Magazine | IEEE Xplore." Accessed: Aug. 15, 2024. [Online]. Available: https://ieeexplore-ieee-org.tudelft.idm.oclc.org/abstract/document/9640612

[62] B. Tan and A. Matta, "The digital twin synchronization problem: Framework, formulations, and analysis," *IISE Transactions*, vol. 56, no. 6, pp. 652–665, Jun. 2024, doi: 10.1080/24725854.2023.2253869.

[63] Y. Jiang *et al.*, "Digital twin-enabled real-time synchronization for planning, scheduling, and execution in precast on-site assembly," *Automation in Construction*, vol. 141, p. 104397, Jun. 2022, doi: 10.1016/j.autcon.2022.104397.

[64] V. Piroumian, "Digital Twins: Universal Interoperability for the Digital Age," *Computer*, vol. 54, no. 1, pp. 61–69, Jan. 2021, doi: 10.1109/MC.2020.3032148.

[65] J. Moyne *et al.*, "A Requirements Driven Digital Twin Framework: Specification and Opportunities," *IEEE Access*, vol. 8, pp. 107781–107801, 2020, doi: 10.1109/ACCESS.2020.3000437.

[66] R. Rezaei, T. K. Chiew, and S. P. Lee, "An interoperability model for ultra large scale systems," *Advances in Engineering Software*, vol. 67, pp. 22–46, Jan. 2014, doi: 10.1016/j.advengsoft.2013.07.003.

[67] A. Tolk and P. Modeling, "Introducing a Reference Model for Measures of Merit for Coalition Interoperability," Jun. 2003.

[68] S. Dittmann, P. Zhang, A. Glodde, and F. Dietrich, "Towards a scalable implementation of digital twins - A generic method to acquire shopfloor data," *Procedia CIRP*, vol. 96, pp. 157–162, Jan. 2021, doi: 10.1016/j.procir.2021.01.069.

[69] "A Scalable Digital Twin Framework Based on a Novel Adaptive Ensemble Surrogate Model | J. Mech. Des. | ASME Digital Collection." Accessed: May 02, 2024. [Online]. Available: https://asmedigitalcollection.asme.org/mechanicaldesign/article/145/2/021701/1148483/A-Scalable-Digital-Twin-Framework-Based-on-a-Novel

[70] K. Alshammari, T. Beach, and Y. Rezgui, "Cybersecurity for digital twins in the built environment: current research and future directions," *ITcon*, vol. 26, pp. 159–173, Apr. 2021, doi: 10.36680/j.itcon.2021.010.

[71] "Cybersecurity in Distributed Power Systems | IEEE Journals & Magazine | IEEE Xplore." Accessed: May 02, 2024. [Online]. Available: https://ieeexplore-ieee-org.tudelft.idm.oclc.org/document/7932449

[72] J. C. Olivares-Rojas, E. Reyes-Archundia, J. A. Gutierrez-Gnecchi, I. Molina-Moreno, J. Cerda-Jacobo, and A. Mendez-Patino, "Towards Cybersecurity of the Smart Grid Using Digital Twins," *IEEE Internet Comput.*, vol. 26, no. 3, pp. 52–57, May 2022, doi: 10.1109/MIC.2021.3063674.

[73] P. Kumar, R. Kumar, A. Aljuhani, D. Javeed, A. Jolfaei, and A. K. M. N. Islam, "Digital twin-driven SDN for smart grid: A deep learning integrated blockchain for cybersecurity," *Solar Energy*, vol. 263, p. 111921, Oct. 2023, doi: 10.1016/j.solener.2023.111921.

[74] I. Yaqoob, K. Salah, M. Uddin, R. Jayaraman, M. Omar, and M. Imran, "Blockchain for Digital Twins: Recent Advances and Future Research Challenges," *IEEE Network*, vol. 34, no. 5, pp. 290–298, Sep. 2020, doi: 10.1109/MNET.001.1900661.

[75] B. Putz, M. Dietz, P. Empl, and G. Pernul, "EtherTwin: Blockchain-based Secure Digital Twin Information Management," *Information Processing & Management*, vol. 58, no. 1, p. 102425, Jan. 2021, doi: 10.1016/j.ipm.2020.102425.

[76] "A Blockchain-Based Approach for the Creation of Digital Twins | IEEE Journals & Magazine | IEEE Xplore." Accessed: May 02, 2024. [Online]. Available: https://ieeexplore-ieee-org.tudelft.idm.oclc.org/document/9001017

[77] S. Huang, G. Wang, Y. Yan, and X. Fang, "Blockchain-based data management for digital twin of product," *Journal of Manufacturing Systems*, vol. 54, pp. 361–371, Jan. 2020, doi: 10.1016/j.jmsy.2020.01.009.

[78] "Federated learning: privacy-preserving data analysis," tno.nl/en. Accessed: Jul. 29, 2024. [Online]. Available: https://www.tno.nl/en/technology-science/technologies/federated-learning/

[79] C. Zhang, S. R. Kuppannagari, R. Kannan, and V. K. Prasanna, "Generative Adversarial Network for Synthetic Time Series Data Generation in Smart Grids," in *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, Oct. 2018, pp. 1–6. doi: 10.1109/SmartGridComm.2018.8587464.

[80] Y. Liu, H. Xie, A. Presekal, A. Stefanov, and P. Palensky, "A GNN-Based Generative Model for Generating Synthetic Cyber-Physical Power System Topology," *IEEE Transactions on Smart Grid*, vol. 14, no. 6, pp. 4968–4971, Nov. 2023, doi: 10.1109/TSG.2023.3304134.

[81] "Energies | Free Full-Text | IoT-Based Digital Twin for Energy Cyber-Physical Systems: Design and Implementation." Accessed: May 02, 2024. [Online]. Available: https://www.mdpi.com/1996-1073/13/18/4762

[82] L. Zhao, I. Brandao Machado Matsuo, Y. Zhou, and W.-J. Lee, "Design of an Industrial IoT-Based Monitoring System for Power Substations," *IEEE Transactions on Industry Applications*, vol. 55, no. 6, pp. 5666–5674, Nov. 2019, doi: 10.1109/TIA.2019.2940668.

[83] H. K. Bundele and K. C. O. Reddy, "IoT Based Phasor Measurement Unit (PMU)," vol. 6, no. 6, 2019.

[84] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge Computing: Vision and Challenges," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637–646, Oct. 2016, doi: 10.1109/JIOT.2016.2579198.

[85] J. Pushpa and S. A. Kalyani, "Chapter Three - Using fog computing/edge computing to leverage Digital Twin," in *Advances in Computers*, vol. 117, 1 vols., P. Raj and P. Evangeline, Eds., in The Digital Twin Paradigm for Smarter Systems and Environments: The Industry Use Cases, vol. 117. , Elsevier, 2020, pp. 51–77. doi: 10.1016/bs.adcom.2019.09.003.

[86] J. Wang, Y. Liu, S. Ren, C. Wang, and S. Ma, "Edge computing-based real-time scheduling for digital twin flexible job shop with variable time window," *Robotics and Computer-Integrated Manufacturing*, vol. 79, p. 102435, Feb. 2023, doi: 10.1016/j.rcim.2022.102435.

[87] C. L. Stergiou and K. E. Psannis, "Digital twin intelligent system for industrial internet of things-based big data management and analysis in cloud environments," *Virtual Reality & Intelligent Hardware*, vol. 4, no. 4, pp. 279–291, Aug. 2022, doi: 10.1016/j.vrih.2022.05.003.

[88] "The Role of AI, Machine Learning, and Big Data in Digital Twinning: A Systematic Literature Review, Challenges, and Opportunities | IEEE Journals & Magazine | IEEE Xplore." Accessed: May 02, 2024. [Online]. Available: https://ieeexplore-ieee-org.tudelft.idm.oclc.org/document/9359733

[89] "Deep Learning in Smart Grid Technology: A Review of Recent Advancements and Future Prospects | IEEE Journals & Magazine | IEEE Xplore." Accessed: Aug. 06, 2024. [Online]. Available: https://ieeexplore-ieee-org.tudelft.idm.oclc.org/abstract/document/9395437

[90] S. J. Russell, P. Norvig, and E. Davis, *Artificial intelligence: a modern approach*, 3rd ed. in Prentice Hall series in artificial intelligence. Upper Saddle River: Prentice Hall, 2010.

[91] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 2, no. 1, pp. 37–52, Aug. 1987, doi: 10.1016/0169-7439(87)80084-9.

[92] A. Tharwat, T. Gaber, A. Ibrahim, and A. E. Hassanien, "Linear discriminant analysis: A detailed tutorial," *AI Communications*, vol. 30, no. 2, pp. 169–190, Jan. 2017, doi: 10.3233/AIC-170729.

[93] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognition*, vol. 77, pp. 354–377, May 2018, doi: 10.1016/j.patcog.2017.10.013.

[94] B. Charbuty and A. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, Art. no. 01, Mar. 2021, doi: 10.38094/jastt20165.

[95] "Support vector machines | IEEE Journals & Magazine | IEEE Xplore." Accessed: Aug. 06, 2024. [Online]. Available: https://ieeexplore-ieee-org.tudelft.idm.oclc.org/abstract/document/708428

[96] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A Search Space Odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222–2232, Oct. 2017, doi: 10.1109/TNNLS.2016.2582924.

[97] "Parameter tuning for dynamic digital twins in inverter-dominated distribution grid - Song - 2020 - IET Renewable Power Generation - Wiley Online Library." Accessed: May 02, 2024. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/full/10.1049/iet-rpg.2019.0163

[98] Y. Golizadeh Akhlaghi *et al.*, "A constraint multi-objective evolutionary optimization of a state-of-the-art dew point cooler using digital twins," *Energy Conversion and Management*, vol. 211, p. 112772, May 2020, doi: 10.1016/j.enconman.2020.112772.

[99] F. Tao, M. Zhang, Y. Liu, and A. Y. C. Nee, "Digital twin driven prognostics and health management for complex equipment," *CIRP*

Annals, vol. 67, no. 1, pp. 169–172, Jan. 2018, doi: 10.1016/j.cirp.2018.04.055.
[100] V. K. Singh and M. Govindarasu, "Decision Tree Based Anomaly Detection for Remedial Action Scheme in Smart Grid using PMU Data," in *2018 IEEE Power & Energy Society General Meeting (PESGM)*, Aug. 2018, pp. 1–5. doi: 10.1109/PESGM.2018.8586159.
[101] A. Saad, S. Faddel, and O. Mohammed, "IoT-Based Digital Twin for Energy Cyber-Physical Systems: Design and Implementation," *Energies*, vol. 13, no. 18, Art. no. 18, Jan. 2020, doi: 10.3390/en13184762.
[102] X. Liu *et al.*, "A systematic review of digital twin about physical entities, virtual models, twin data, and applications," *Advanced Engineering Informatics*, vol. 55, p. 101876, Jan. 2023, doi: 10.1016/j.aei.2023.101876.
[103] "Data Platform - Data Federation," Oracle Help Center. Accessed: Jul. 31, 2024. [Online]. Available: https://docs.oracle.com/en/solutions/data-platform-federation/index.html
[104] E. Y. Song, M. Burns, A. Pandey, and T. Roth, "IEEE 1451 Smart Sensor Digital Twin Federation for IoT/CPS Research," in *2019 IEEE Sensors Applications Symposium (SAS)*, Mar. 2019, pp. 1–6. doi: 10.1109/SAS.2019.8706111.
[105] J. Zhang *et al.*, "BioMart: a data federation framework for large collaborative projects," *Database (Oxford)*, vol. 2011, p. bar038, Sep. 2011, doi: 10.1093/database/bar038.
[106] M. Singh, E. Fuenmayor, E. P. Hinchy, Y. Qiao, N. Murray, and D. Devine, "Digital Twin: Origin to Future," *Applied System Innovation*, vol. 4, no. 2, Art. no. 2, Jun. 2021, doi: 10.3390/asi4020036.
[107] A. Dognini *et al.*, "Blueprint of the Common European Energy Data Space," Zenodo, Apr. 2024. doi: 10.5281/zenodo.10964387.
[108] *Power System Modelling and Scripting*. Accessed: May 02, 2024. [Online]. Available: https://link.springer.com/book/10.1007/978-3-642-13669-6
[109] M. Manukyan and G. Gevorgyan, "Canonical Data Model for Data Warehouse," Aug. 2016, pp. 72–79. doi: 10.1007/978-3-319-44066-8_8.
[110] S. Auer, V. Kovtun, M. Prinz, A. Kasprzik, M. Stocker, and M. E. Vidal, "Towards a Knowledge Graph for Science," in *Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics*, Novi Sad Serbia: ACM, Jun. 2018, pp. 1–6. doi: 10.1145/3227609.3227689.
[111] Y.-M. Baek, J. Song, Y.-J. Shin, S. Park, and D.-H. Bae, "A meta-model for representing system-of-systems ontologies," in *Proceedings of the 6th International Workshop on Software Engineering for Systems-of-Systems*, Gothenburg Sweden: ACM, May 2018, pp. 1–7. doi: 10.1145/3194754.3194755.
[112] M.-S. Baek, "Digital Twin Federation and Data Validation Method," in *2022 27th Asia Pacific Conference on Communications (APCC)*, Oct. 2022, pp. 445–446. doi: 10.1109/APCC55198.2022.9943622.
[113] "A systematic overview of data federation systems | www.semantic-web-journal.net." Accessed: Feb. 27, 2024. [Online]. Available: https://www.semantic-web-journal.net/content/systematic-overview-data-federation-systems
[114] J. Conde, A. Munoz-Arcentales, Á. Alonso, S. López-Pernas, and J. Salvachúa, "Modeling Digital Twin Data and Architecture: A Building Guide With FIWARE as Enabling Technology," *IEEE Internet Computing*, vol. 26, no. 3, pp. 7–14, May 2022, doi: 10.1109/MIC.2021.3056923.
[115] "Java JDBC API." Accessed: Jun. 05, 2024. [Online]. Available: https://docs.oracle.com/javase/8/docs/technotes/guides/jdbc/
[116] David-Engel, "ODBC API Reference - ODBC API Reference." Accessed: May 02, 2024. [Online]. Available: https://learn.microsoft.com/en-us/sql/odbc/reference/syntax/odbc-api-reference?view=sql-server-ver16
[117] S. Jimenez, "Interoperability Framework in Energy Data Spaces," Zenodo, Oct. 2023. doi: 10.5281/zenodo.10117882.
[118] "What Is Data Modeling? | IBM." Accessed: May 02, 2024. [Online]. Available: https://www.ibm.com/topics/data-modeling
[119] P. P.-S. Chen, "The entity-relationship model—toward a unified view of data," *ACM Trans. Database Syst.*, vol. 1, no. 1, pp. 9–36, Mar. 1976, doi: 10.1145/320434.320440.
[120] R. J. Muller, *Database Design for Smarties: Using UML for Data Modeling*. Morgan Kaufmann, 1999.
[121] S. M. Rashid *et al.*, "The Semantic Data Dictionary - An Approach for Describing and Annotating Data," *Data Intell*, vol. 2, no. 4, pp. 443–486, 2020, doi: 10.1162/dint_a_00058.
[122] "SAREF: the Smart Applications REFerence ontology." Accessed: May 02, 2024. [Online]. Available: https://saref.etsi.org/core/v3.2.1/
[123] R. van der Weerdt, V. de Boer, L. Daniele, and B. Nouwt, "Validating SAREF in a Smart Home Environment," in *Metadata and Semantic Research*, E. Garoufallou and M.-A. Ovalle-Perandones, Eds., Cham: Springer International Publishing, 2021, pp. 35–46. doi: 10.1007/978-3-030-71903-6_4.
[124] "CEN-CENELEC-ETSI Smart Grid Coordination Group".
[125] R. Santodomingo *et al.*, "SGAM-based methodology to analyse Smart Grid solutions in DISCERN European research project," in *2014 IEEE International Energy Conference (ENERGYCON)*, May 2014, pp. 751–758. doi: 10.1109/ENERGYCON.2014.6850510.
[126] "Home - Gaia-X: A Federated Secure Data Infrastructure." Accessed: May 02, 2024. [Online]. Available: https://gaia-x.eu/
[127] B.-B. für W. und Klimaschutz, "Das Gaia-X Ökosystem -." Accessed: May 02, 2024. [Online]. Available: https://www.bmwk.de/Redaktion/DE/Dossier/gaia-x.html
[128] S. Forte, F. Gast, T. Dickopf, C. Sturmer, and M. Weber, "Gaia-X based Data Ecosystem Offerings in the Context of Data Management Systems," in *2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation (ETFA)*, Sep. 2024, pp. 1–6. doi: 10.1109/ETFA61755.2024.10710761.
[129] "Common Information Model (CIM) for Grid Models Exchange." Accessed: Jun. 05, 2024. [Online]. Available: https://www.entsoe.eu/digital/common-information-model/cim-for-grid-models-exchange/
[130] "CIM | DMTF." Accessed: May 02, 2024. [Online]. Available: https://www.dmtf.org/standards/cim
[131] "Common Information Model (CIM)." Accessed: May 02, 2024. [Online]. Available: https://www.entsoe.eu/digital/common-information-model/
[132] "IEC TR 63097:2017 | IEC Webstore | smart grid, smart energy." Accessed: Jun. 05, 2024. [Online]. Available: https://webstore.iec.ch/publication/27785
[133] A. Bytyqi, S. Gandhi, E. Lambert, and N. Petrovic, "A Review on TSO-DSO Data Exchange, CIM Extensions and Interoperability Aspects," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 2, pp. 309–315, 2022, doi: 10.35833/MPCE.2021.000770.
[134] J. Fremont, E. Lambert, C. Bouquet, O. Carre, D. Ilhat, and P. Metayer, "CIM extensions for ERDF information system projects," in *2009 IEEE Power & Energy Society General Meeting*, Jul. 2009, pp. 1–5. doi: 10.1109/PES.2009.5275789.
[135] "EU-SysFlex." Accessed: May 02, 2024. [Online]. Available: https://eu-sysflex.com/
[136] "TDX-ASSIST - Home." Accessed: May 02, 2024. [Online]. Available: http://www.tdx-assist.eu/index.php
[137] "Homepage," OneNet Project. Accessed: May 02, 2024. [Online]. Available: https://onenet-project.eu/
[138] A. Naumann, I. Bielchev, N. Voropai, and Z. Styczynski, "Smart grid automation using IEC 61850 and CIM standards," *Control Engineering Practice*, vol. 25, pp. 102–111, Apr. 2014, doi: 10.1016/j.conengprac.2013.12.001.
[139] Q. Xia, W. Ye, Z. Tao, J. Wu, and Q. Li, "A survey of federated learning for edge computing: Research problems and solutions," *High-Confidence Computing*, vol. 1, no. 1, p. 100008, Jun. 2021, doi: 10.1016/j.hcc.2021.100008.
[140] H. T. Reda *et al.*, "Vulnerability and Impact Analysis of the IEC 61850 GOOSE Protocol in the Smart Grid," *Sensors (Basel)*, vol. 21, no. 4, p. 1554, Feb. 2021, doi: 10.3390/s21041554.
[141] R. Ala-Laurinaho, J. Autiosalo, A. Nikander, J. Mattila, and K. Tammi, "Data Link for the Creation of Digital Twins," *IEEE Access*, vol. 8, pp. 228675–228684, 2020, doi: 10.1109/ACCESS.2020.3045856.
[142] M. Dietz, B. Putz, and G. Pernul, "A Distributed Ledger Approach to Digital Twin Secure Data Sharing," in *Data and Applications Security and Privacy XXXIII*, S. N. Foley, Ed., Cham: Springer International Publishing, 2019, pp. 281–300. doi: 10.1007/978-3-030-22479-0_15.

MAY MYAT THWE (Student Member, IEEE) received the B.Sc. (Honors) degree in Computer Science from University of Yangon in 2018 and M.S. degree in Computer Science from Korea Advanced Institute of Science and Technology (KAIST) in 2022. She is currently a Doctoral Researcher in the Intelligent Electrical Power Grids group at the Department of Electrical Sustainable Energy, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology. Her current research interests include digital twin, machine learning, and cybersecurity of power grids.

Alexandru Ştefanov (Member, IEEE) received the M.Sc. degree from the University Politehnica of Bucharest, Romania, in 2011, and the Ph.D. degree from University College Dublin, Ireland, in 2015. He is Associate Professor of Intelligent Electrical Power Grids at TU Delft, The Netherlands. He is the Director of the Control Room of the Future (CRoF) Technology Centre. He is leading the Cyber Resilient Power Grids (CRPG) research group. His research interests include cybersecurity of power grids, resilience of cyber-physical systems, and next generation grid operation. He holds the Professional Title of Chartered Engineer from Engineers Ireland.

Vetrivel Subramaniam Rajkumar (Student Member, IEEE) received the bachelor’s degree in electrical engineering from Anna University, India, in 2013, and the M.Sc. degree in electrical power engineering from the Delft University of Technology, The Netherlands, in 2019, where he is currently a Doctoral Researcher with the Intelligent Electrical Power Grids Group, Department of Electrical Sustainable Technology. His current research interests include cybersecurity for power grids and impact analysis of cyber-attacks on power systems.

Peter Palensky (Senior Member, IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. and Habilitation degrees from the Vienna University of Technology, Austria, in 1997, 2001, and 2015, respectively. He co-founded Envidatec, a German startup on energy management and analytics. In 2008, he joined the Lawrence Berkeley National Laboratory, Berkeley, CA, USA, as a Researcher, and the University of Pretoria, South Africa. In 2009, he became the Head of the Business Unit, Austrian Institute of Technology in Sustainable Building Technologies, where he was the first Principal Scientist of Complex Energy Systems. In 2014, he was appointed as a Full Professor of Intelligent Electric Power Grids with TU Delft, The Netherlands. He is active in international committees, such as ISO or CEN. His research interests include energy automation networks, smart grids, and modeling intelligent energy systems. He also serves as an IEEE IES AdCom Member-at-Large in various functions for IEEE. He is the past Editor-in-Chief of IEEE Industrial Electronics Magazine and an Associate Editor of several other IEEE publications and regularly organizes IEEE conferences.