Highlights

Survey and Insights on Digital Twins Design and Smart Grid's Applications

Sonia Djebali, Guillaume Guerard, Ihab Taleb

- Literature review about Digital Twin's Design and Modeling Process.
- Description of a digital twin management design called DTOps.
- Citation of the most recent and relevant surveys and works in the domain of Digital Twins for Smart Grid applications.

Survey and Insights on Digital Twins Design and Smart Grid's Applications

Sonia Djebali[*,a], Guillaume Guerard[a], Ihab Taleb[a]

[a]*Léonard de Vinci Pôle Universitaire, Research Center, Paris La Défense, 92 916, France*

Abstract

Digital twins are a promising technology for simulating complex systems, especially in the smart grid domain. This paper offers a comprehensive literature review on digital twins, focusing on data gathering, data management, and human-in-the-loop control design aspects. Emphasizing the integration of AI and machine learning in big data, it enhances analytics and decision-making capabilities. We introduce a collaborative framework involving multiple stakeholders to maximize the potential of digital twins. The paper examines digital twin applications in smart grids, covering areas like asset management, predictive maintenance, energy optimization, and demand response. By synthesizing research and implementation findings, we identify trends, challenges, and opportunities in the field.

Keywords: Digital Twin, DevOps, Smart Grid, Design Methodology

1. Introduction

Digital Twin (DT) technologies have emerged as a transformative concept in the context of Smart Grid (SG) applications, revolutionizing the way we monitor, model, and control power systems. The definition of DT, as summarized by Fuller et al. (2020), entails a virtual replica of a physical system or process that mimics its behavior in real-time, providing valuable insights and

[*]Corresponding author

Email addresses: sonia.djebali@devinci.fr (Sonia Djebali), guillaume.guerard@devinci.fr (Guillaume Guerard), ihab.taleb@devinci.fr (Ihab Taleb)

facilitating decision-making processes. In line with the SG's challenges identified by Amin and Wollenberg (2005); Amin (2011), DT technologies offer immense potential for optimizing energy efficiency, improving grid reliability, and fostering sustainability in SGs.

DT is not limited to SG applications. For example, DTs are used in production lines to optimize processes and manage energy, heat, time lag, and overall architecture, as demonstrated by Min et al. (2019) in the petrochemical industry and by Mendi (2022) in the automobile industry. As presented by Tao et al. (2019), DTs are closely linked to Industry 4.0. In this paradigm, cyber-physical systems (CPS) and DTs have garnered extensive attention from researchers and industry practitioners. The boundary between CPSs and DTs is subtle, and discussions to define each are actively ongoing. In this article, we consider DTs as digital representations of physical systems that may not yet be efficient, with the aim of improving their effectiveness.

The effective design of DT systems is crucial for their development and implementation in SGs. It involves considerations such as system architecture, data acquisition from heterogeneous sources, scalability, and replicability. Given the complexity of SG systems, as evidenced by Guérard et al. (2012); Ahat et al. (2013), the design process ensures an accurate representation of all system components and allows for the seamless integration of various SG technologies and components (which potentially don't yet exist).

By representing and organizing knowledge in a structured manner, ontologies enable semantic interoperability, data integration, and meaningful contextual understanding. De Nicola and Villani (2021) declare ontologies capture the complex relationships and dependencies among elements to provide decision-making processes accurately.

Data management is another critical aspect of DT systems as highlighted by Daki et al. (2017). SGs generate vast amounts of data and its impact mustn't be neglected (Allam and Dhunny (2019)). Data come from diverse sources, including sensors, SCADA systems, and IoT devices. Effective data management techniques, such as data acquisition, preprocessing, fusion, and quality assurance, ensure the reliability and integrity of the DT. Additionally, data storage, analytics, and visualization, as part of the overall design, enable valuable insights and support data-driven decision-making processes.

From ontologies and data management, one can model and simulate SGs. Accurate system modeling allows the DT to mirror the behavior of the physical grid, enabling real-time monitoring, analysis, and prediction of grid dynamics. Different modeling approaches, such as physics-based models, data-driven models, and hybrid models, capture the intricacies of the SG's operation. Simulation and validation techniques further enhance the reliability and accuracy of DT systems, allowing for virtual testing, scenario analysis, and performance evaluation.

While DT technologies offer significant opportunities for SG applications, they also come with challenges. Data quality, integration, and interoperability pose obstacles to harnessing the full potential of DT systems (Jafari et al. (2023)). Ensuring privacy and security of grid data in a DT environment is also a critical concern (Sakhnini et al. (2021)). Additionally, the computational requirements for large-scale DT systems and the need for standardization present challenges that need to be addressed.

Many studies propose a design methodology for handling DTs. Psarommatis and May (2023) provided a design after analyzing 760 papers; however, their method is generic and lacks details about the systemic approach, development, and operations. On the other hand, Tao et al. (2019) introduced a V-method with plenty of local feedback but lacks an internal loop if a previous step doesn't correspond to the needs. Schroeder et al. (2020) encompassed human-machine interaction and learning processes but failed to provide a complete design. In this context, our paper fills the gap concerning a methodology design for DTs. Moreover, we discuss the role of data management and technology to maintain a dynamic workflow and ensure a relevant design for complex system modeling.

To summarize, we have identified the following gaps in DT methodology design:

- The lack of commonly used standardized frameworks and protocols as exposed by Jacoby and Usländer (2020). Ensuring that DTs from different domains, industries, and software platforms can seamlessly interact remains a significant challenge.
- DTs rely heavily on data as highlighted by Singh et al. (2021b). Data flow management must be a part of the DT design.
- DTs are mostly done for specific use-cases (Attaran and Celik (2023)). DTs are being applied across various domains, from manufacturing to healthcare to urban planning. Data management, software and design are differents with little overlap between studies.

- Security, ethic and privacy are often neglected (Farsi et al. (2020)). As DTs collect and process sensitive data, security must be a part of the DT design.
- Human-Machine interaction due to the use of machine learning, AI integration, robot and cobotic is an issue to how to manage a DT as exposed by Wang et al. (2022). The interaction must be an important part of such design.

The article contributions are as follows:

- A literature review on SGs' DT applications with up-to-date references
- An updated DT definition taking account of recent works and technologies
- A methodology to design DT, called DTOps with 4 knots, adapted to recent works in large-scale software development methodologies
- The knot about Data is discussed, especially how to manage ontologies and data management in such methodology.
- A discussion about main challenges on various use cases and how the presented methodology can solve those.

The objective of this article is to delve into the various components for developing efficient and effective DT systems in the context of SG applications. By exploring and analyzing the design considerations (Section 3), ontologies (Section 4), data management (Section 4), and modeling and simulation approaches (Section 5), we aim to provide valuable insights into the integration of these components to enable the realization of advanced DT systems for SGs. Then, some applications (Section 5) are analyzed. Challenges and future directions are discussed in Section 6. The section 7 concludes the paper.

2. Definition of Digital Twin

A DT can be defined as a virtual replica or digital representation of a physical asset, process, or system. It can be defined as follows:

> A DT refers to a virtual representation or digital replica of a physical object, system, or process. It encompasses both the physical and digital realms, allowing real-time synchronization, interaction, and feedback between the physical entity and its digital counterpart. DTs enable monitoring, analysis, simulation, and optimization, facilitating enhanced understanding, decision-making, and performance improvements in various domains, including manufacturing, infrastructure, and the Internet of Things (IoT).

2.1. Key Components

The key components of a DT system encompass various aspects, each playing a vital role in capturing and replicating reality in the virtual space. The works of Jones et al. (2020); Singh et al. (2021a); Tao et al. (2022) define the following key components to build a DT:

Physical Object/System. The physical object or system is the real-world entity that the DT represents. It can be a physical asset, such as a machine, building, or infrastructure system, or it can represent a larger system, such as a smart city or a manufacturing plant.

Virtual Representation. The virtual representation is the digital counterpart of the physical object or system. It is created using various technologies, such as computer-aided design, 3D modeling, or point cloud data. The virtual representation mirrors the geometry, structure, and attributes of the physical entity.

Data Acquisition. It rely on the collection of vast amounts of data from diverse sources such as sensors, meters, SCADA systems, and IoT devices. These data streams provide real-time and historical information about operating conditions and performance.

Data Integration and Fusion. A DT must incorporate methods to integrate and fuse complex data (numeric, qualitative, text, time series) from heterogeneous sources and/or providers. This process involves harmonizing and aggregating data to create a unified view of the SG, enabling a comprehensive understanding of its behavior.

Modeling and Simulation. DT systems utilize advanced modeling techniques to replicate the physical behavior of the SG. These models can encompass various levels of complexity, ranging from physics-based models that capture the fundamental principles governing the grid's operation to data-driven models that leverage historical data to make predictions and optimize energy management.

Visualization and User Interface. It is a crucial component of DT systems as it allows operators and stakeholders to gain insights and make informed decisions. Intuitive user interfaces and visual representations enable stakeholders to monitor the grid's status, identify anomalies, and assess the impact of different operational scenarios.

2.2. Digital Twin Levels

The role of DT systems in SGs is multifaceted. They provide a comprehensive and dynamic representation of the physical grid, allowing operators and stakeholders to gain a deeper understanding of the system's behavior, diagnose problems, and optimize operations. As defined by Amor et al. (2019), SG's modeling and simulation enable real-time monitoring, providing users with a holistic view of the grid's status, helping them identify anomalies or potential risks promptly. Moreover, by simulating various operational scenarios and performing predictive analytics, a model supports decision-making processes, enabling proactive interventions to improve energy efficiency, reliability, and grid stability. We can summarize the degree of maturity of DTs into 6 additive levels :

1. Foundation Model: Virtual model of a physical object for asset information management, real-time data collection on a cloud/on-premise platform.
2. Predictive DT: This model adds real-time analytics for predictive maintenance and performance management.
3. Prescriptive DT: This model enhances the previous ones with real-time optimization and what-if simulation.
4. Transformative DT: Remote Collaboration and immersive training are the key points of this level;
5. Cognitive DT: This level adds machine-human interactions as active learning to let more place to artificial-based decision-making, automatic maintenance, and autonomous operations (through drones for

example). See the surveys of Ramu et al. (2022); Zheng et al. (2022) for more details on this level.

6. Adaptive Autonomous Twin: It can also change their learning strategy to adapt to new environmental uncertainties. Hribernik et al. (2021) provide more explanation of this concept.

DTs in industry predominantly fall within the range of levels 2 or 3, with a few large-scale DTs reaching level 4. Now, let's delve into some important use cases to illustrate this.

2.3. Digital Twin Applications

A DT can be applied to various use cases, which we can categorize into four main concepts: Monitoring and Control, Maintenance, Decision Support, and Training. Let's explore each of these concepts individually. It's important to note that a DT can encompass one or multiple of these use cases.

Monitoring and Control. DTs facilitate remote monitoring and control of the physical system. They allow access and control of the virtual representation of the system from anywhere, providing remote visibility, diagnostics, and control capabilities. The use of dashboards or online boards about our consumption become common nowadays.

Real-time monitoring and control of the SG is a key point of a DT. By continuously updating the virtual representation based on real-time data, operators can closely observe a system's performance, identify potential issues, and take proactive measures with, for example, Energy Management Systems. For example, Francisco et al. (2020) propose a DT to efficiently manage the building's energy in real-time. DTs-based remote semi-physical commissioning of flow-type smart manufacturing systems is also an example of how this technology can be applied. By creating digital replicas of manufacturing systems, it becomes possible to commission and fine-tune systems remotely, saving both time and resources.

DTs support the optimization and control of the physical system by applying algorithms and models to identify optimal operating parameters, configuration settings, or resource allocation. This functionality enables optimizing energy consumption, improving efficiency, and maximizing performance. For example, Bhatti et al. (2021) did a survey about the use of DTs for electric vehicles and how they can be integrated into the grid for better optimization and control of energy flows.

Maintenance. By analyzing real-time and historical data, DTs predict potential failures or performance degradation in the physical system. This enables proactive maintenance, helping to prevent downtime, reduce costs, and extend the lifespan of the assets. Building Information models (BIM)s are recent examples of this goal. For example, Khajavi et al. (2019) provides a review of DT benefits for building, thanks to building information modeling, and how predictive maintenance is essential in those models.

DTs cover the entire lifecycle of the physical system, from design and construction to operation and maintenance. They enable the capture and integration of data at each stage, supporting informed decision-making, performance evaluation, and optimization throughout the lifecycle. As an example, Yitmen et al. (2021) present a Cognitive DT model for building lifecycle management.

Decision Support. DTs provide decision-makers with a virtual platform to test different scenarios, evaluate the impact of their decisions, and identify optimal strategies. By incorporating human expertise and judgment, the simulation outcomes can be refined, leading to better-informed decisions. In dynamic environments where changes are frequent, DTs can facilitate quick adjustments and optimizations, ensuring that systems remain aligned with changing needs and outcomes.

DTs enable simulation and what-if analysis, allowing stakeholders to simulate different scenarios and assess their potential impact on the physical system. This functionality helps in testing hypotheses, evaluating alternative strategies, and optimizing decision-making. By capturing data and insights from the physical system, one can identify areas for improvement, test new technologies or processes, and drive innovation in the design, operation, and management of the system. For example, Taleb. et al. (2023) propose a multi-agent model to simulate various scenarios of grid failures or enhancement in the context of an island.

By analyzing the simulation outcomes, operators can identify patterns, trends, and opportunities for optimizing system performance, leading to ongoing refinement of operational strategies. In emergency situations, such as blackouts, equipment failures, or natural disasters, the result of such simulation helps operators to develop and refine emergency response plans. It enables them to understand the potential consequences of their decisions, assess the effectiveness of different response strategies, and identify areas for improvement. DTs allow for the validation of system performance in a semi-physical simulation manner. Stakeholders can thoroughly test system configurations and changes before implementing them in the physical environment. This reduces the likelihood of costly errors and downtime.

A DT has the potential to facilitate efficient configuration/reconfiguration, validation, and testing processes in various systems. It enables the swift identification of inefficiencies and allows for the assessment of the viability of physical solutions during implementation. For instance, DTs play a pivotal role in Industry 4.0 by supporting the design of smart manufacturing systems, aiding in the rapid customization of automated flow-shop manufacturing systems, and assisting in the creation of comprehensive models for configuring, controlling, optimizing, and managing flow-type smart manufacturing systems (Tao et al. (2022)).

Training. DT provides a platform for collaborative decision-making, allowing stakeholders to actively participate, contribute their insights, and understand the implications of different grid management strategies.

DT allows them to gain experience and improve their skills in managing complex grid operations. They can practice responding to different scenarios, learn from their mistakes, and develop strategies for efficient and effective grid management. This functionality helps in enhancing skills, evaluating strategies, and improving safety without impacting the physical system. For example, Olszewski et al. (2019) propose a serious game for the ecological impact of people in a smart city.

3. DTOps Methodology for DT

3.1. Needs for an Efficient DT

Designing DTs for complex adaptive systems, such as SGs, smart cities, or smart buildings, requires a thoughtful and systematic approach. These systems exhibit dynamic and interconnected behaviors, where numerous components interact and adapt in response to changing conditions. To effectively design DTs, specific methods and considerations are needed:

- System Understanding and Mapping
- Modularity and Scalability
- Data Acquisition and Integration

- Dynamic Modeling and Simulation
- Feedback Loops and Control Mechanisms
- Resilience and Adaptability
- Visualization and User Interface

System Understanding and Mapping. Any design begins with a deep understanding of the system's structure, components, and interactions. This involves mapping out the various elements, subsystems, and their interdependencies. For SGs, this may include power generation, distribution networks, energy storage, demand response systems, and renewable energy integration. As an example, Amor et al. (2014) propose a multi-level model of a SG by analyzing each component of the grid.

Modularity and Scalability. One should consider a modular approach, where components can be added, removed, or updated without disrupting the entire system. This allows for flexibility and adaptability as the system evolves or expands. For example, Egert et al. (2021) propose an holonic multi-agent model for SG where each component adapts to new conditions and what-if scenarios.

Data Acquisition and Integration. Design methods should address data acquisition, considering the integration of heterogeneous data sources, such as sensors, IoT devices, utility networks, and external data feeds. Ensuring data compatibility, reliability, and real-time accessibility is crucial. Zagan and Danubianu (2020) present the data lake approach and how it can be used efficiently in the context of SGs.

Dynamic Modeling and Simulation. Any design should incorporate dynamic modeling techniques, including agent-based modeling, system dynamics, or other simulation approaches. These models capture the system's adaptive behavior, allowing for scenario testing and prediction of system dynamics. For example, Guerard and Pousseur (2020) present a JADE model for smart cities to simulate how consumers in a microgrid must adapt to grid defects.

Feedback Loops and Control Mechanisms. A model has to incorporate feedback mechanisms to update the virtual representation based on real-time data from the physical system. This enables closed-loop control, adaptive decision-making, and system optimization. Since SG integrates socio-economic-technical elements, feedback helps the process to adapt to conditions and constraints as shown by the increasing amount of dynamic methods for SGs like the works of Mbungu et al. (2020).

Resilience and Adaptability. Resilience features, such as fault tolerance, redundancy, and adaptive control strategies are part of a complex system. This allows the DT to mirror the system's ability to adapt and recover from disturbances. Resilience occurs at multiple levels of scale and time and on different actors of the system as presented by Kandaperumal and Srivastava (2020).

Visualization and User Interface. Design methods should include visualization techniques and user-friendly interfaces to facilitate interaction with the DT. This enables monitoring system behavior, access to relevant information, and make informed decisions. Visualization tools can include 3D models (like the Smart City construction and management by Lv et al. (2022)), dashboards (like Helibot by Nguyen et al. (2020) to manage peer-to-peer energy's exchange), and augmented reality interfaces (Lodetti et al. (2021) propose to help electricians thanks to VR).

3.2. DevOps Methodology

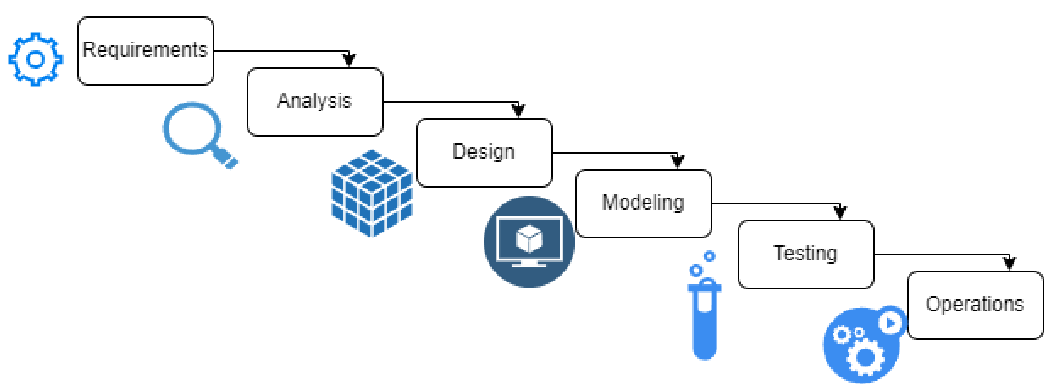

Figure 1: Waterfall modeling methodology.

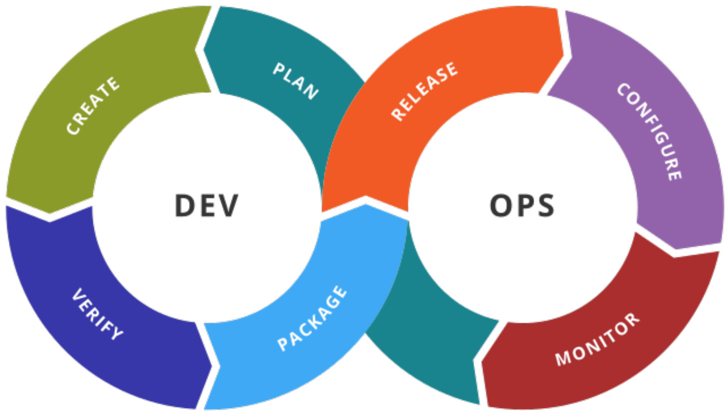

Figure 2: DevOps methodology. Source Wikimedia Commons.

A classic Waterfall methodology as explained in Figure 1 can not handle the complexity of such design. However, the DevOps methodology (shown in Figure 2) shares some similarities with the design of DTs for complex systems (see the works of Leite et al. (2019)). DevOps is an iterative and collaborative approach that focuses on continuous improvement, adaptability, and scalability in software development and operations. The DevOps methodology introduces several new paradigms, including:

- Continuous Integration and Deployment CI/CD
- Colloboration and Communication
- Continuous Improvement
- Human-in-the-loop system

Continuous Integration and Deployment. Regular data integration and real-time updates from the physical system are essential for maintaining an accurate representation of the system's behavior. Since the DT can handle What-If scenarios, the model must also include the possibility to add, remove, and alter existing sub-system.

Collaboration and Communication. DT design involves multiple stakeholders, and effective collaboration among domain experts is essential. Designing a DT requires knowledge of many fields and the full cooperation of experts to handle the complexity of the physical system.

Continuous Improvement. Like any process of complex adaptive system modeling, it needs to identify areas for enhancement, address system inefficiencies, and introduce iterative improvements to enhance performance and value delivery.

Human-in-the-loop system. It refers to the integration of human decision-makers, operators, or stakeholders into the modeling and simulation processes, enabling their active participation and influence on system behavior and outcomes. Human-in-the-loop has been integrated into active learning and now impacts several other fields, see Rothrock and Narayanan (2011) for more details. The Figure 3 show the difference between a human at endpoints (on top) and human-in-the-loop human-machine interactions.

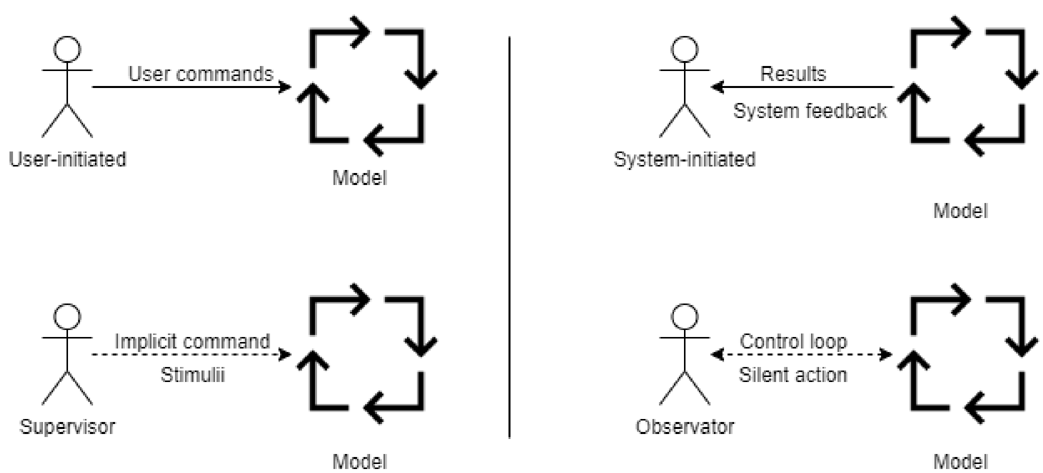

Figure 3: Difference between human-out (top)/human-in (bottom) -the-loop.

A parallel between needs and DevOps methodology provides a better understanding of the necessity to adapt DevOps to DT design.

- System Understanding and Mapping: DevOps enhances system understanding by promoting communication and collaboration among stakeholders at each iteration. This improved knowledge enables better project mapping, accounting for new challenges, gaps, and evolving needs identified throughout the methodology.

- Modularity and Scalability: They are key principles in DevOps, particularly in the context of the DevOps loop and CI/CD. These concepts emphasize the importance of designing systems that can be easily broken down into modular components and scaled horizontally to accommodate changing requirements and workloads.

- Data Acquisition and Integration: It is not typically managed within the DevOps methodology. However, before modeling a physical object, it is crucial to gather relevant data about it. In essence, a separate loop dedicated to Data Acquisition and Integration should be integrated into the methodology's design.

- Dynamic Modeling and Simulation: It requires adaptations when applying the DevOps methodology. To address this, we adapt CI/CD and Continuous Improvement principles to suit our specific case study.

- Feedback Loops and Control Mechanisms: It involves the 'human-in-the-loop' paradigm, where humans provide valuable feedback in addition to machine-generated feedback through control mechanisms. This iterative process aims to continuously improve the model with each iteration..

- Resilience and Adaptability: The modeling process involves the continuous evolution of the complex system at each loop. Some functions may become obsolete, while others require updates. Achieving resilience and adaptability relies on CI/CD and Continuous Improvement principles.

- Visualization and User Interface: They are not inherent parts of the DevOps methodology but typically fall within the operations phase of the design methodology.

By adopting the DevOps methodology, we can sketch a first DT design:

- Real Product:
  1. Gather data;
  2. Integrate data;
  3. Choose Design.
- Digital Twin:

1. Evaluate and analyze data;
2. Model characteristics;
3. Evaluate thanks to simulations behaviors to stimuli change;
4. Transform to feedback.

- Make adjustments, return to first point.

Now that we've introduced the core concepts of the DevOps methodology and their application to the requirements of an efficient DT, we will proceed to describe the main concepts of our methods and the key components in the next subsection.

3.3. DTOps Methodology

To summarize the methodology presented in this paper, DT design should include the following processes and paradigms adapted mainly from DevOps, but also from CRISP-ML(Q) (Studer et al. (2021)), DevSecOps (Akbar et al. (2022)), ML-factory (Lévy et al. (2022)), and MLOps (Kreuzberger et al. (2023)):

1. Define and Design
2. Data Acquisition and Integration
3. Model Development and Training
4. Deployment and Operations

Define and Design. In the initial phase, the DT is defined, and the design considerations are established. This involves identifying the scope, objectives, and desired outcomes of the DT. The domain knowledge and requirements are analyzed, and the DT architecture, data models, and interfaces are designed.

Data Acquisition and Integration. The DT relies on data from various sources. In this phase, data acquisition mechanisms are established to collect data from sensors, devices, and external systems. Data integration techniques are applied to aggregate, preprocess, and cleanse the data, ensuring its quality and consistency.

Model Development and Training. DTs leverage models to simulate and predict system behavior. In this phase, models are developed, ranging from physics-based models to machine learning algorithms. The models are trained using historical data, and techniques such as MLOps are employed for model development, versioning, and deployment.

Deployment and Operations. Once the DT is developed and trained, it is deployed in the operational environment. Real-time data from the physical system is continuously fed into the DT, allowing it to monitor and analyze the system's behavior. DevSecOps principles are applied to ensure secure and reliable deployment, with continuous monitoring, updates, and bug fixes.

The whole methodology can be summarized into a knot lifecycle (Figure 4) where each part forms a loop to itself and can provide feedback (i.e. get back to) any previous part:

1. Design: Define the physical model and apply a systemic approach to decompose it.
   (a) Physical Model: the stakeholders have to define the scope of what they needs to know (as data) from the physical model.
   (b) Systemic Analysis: the stakeholders have to study the physical model thanks to a systemic analysis, see Lévy et al. (2022) for more details about this method.
   (c) Define: since the systemic analysis is done, one has to use a modeling language to formalize the model as shown by Ahat et al. (2013). We recommend using a language adapted to complex systems such as AUML, SysML with KQML or ACL following the FIPA or OMG frameworks (as exposed by Guerard and Pousseur (2020)). Ontologies are created at this step.
   (d) Refine: The Define step is compared to the scope of the Physical Model. If necessary, one can return to ($1.a$) to modify the scope.
2. Data: From the design, work on the data to fulfill requirements. This knot is discussed in the following section.
   (a) Collect: At this step, the data is gathered into Datamesh or Lakehouse. We recommend reading Lyko et al. (2016) for a better understanding of the lifecycle of data from the sensors to the storage.
   (b) Curate: The organization and integration of data collected from various sources follow annotation, publication, presentation, and add values are created from the data. We recommend reading Freitas and Curry (2016) for a better understanding of the goals and processes of Data curation.
   (c) Transform: Governance, orchestration, and lineage of the data mustn't be neglected. We recommend reading the books of Reis and Housley (2022) about those steps.

(d) Validate: Data quality, versioning, and features validate or not this knot. If necessary, one can return to (2.a) or to (1.a) if the data processing doesn't satisfy the proposed design.

3. Model: Integrate the data into the design, define the scenarios and the wanted behaviors.
   (a) Build: Since the model and the data are ready. Programmers have to build the model thanks to an adapted language. We recommend using multi-agent languages such as GAMA, JADE, NetLogo or AnyLogic. Agent-based modeling follows its own frameworks (see Niazi and Hussain (2012))
   (b) Train: Since complex systems are sensible to internal and external factors, monitors, operators and supervisors (respectively machine operations, human operations, and human supervision of the machine operations) have to be set. The sensibility of the model is tested. We recommend seeing Dooley (1997) for the chaos theory and attractors to understand this step.
   (c) Evaluate: Various variables of the model are tracked during the first simulations to evaluate if the program corresponds to the model (entities, behaviors).
   (d) Test: Feedback about the training and evaluation can validate this knot (continue to Operations first step) or jump to (4.d) to provide whole feedback to Design and Data knots.
4. Operations: Test the model, and provide feedback from humans and machines.
   (a) Deploy: The model is tested in real conditions.
   (b) Operate: The Train and Evaluate steps are done with real data and in real conditions.
   (c) Monitor: Supervisors provide feedback about how the model performs in real conditions.
   (d) Plan: Depending on the feedback, stakeholders of each step have to plan how they need to change or update their work.
   (e) Formulate: The plan of change follows its own framework to be relevant and understandable for all stakeholders. Return to each step (a) for each knot.

Similar to the iterative nature of DT design, DTOps emphasizes iterative development and improvement cycles. It promotes a feedback-driven process

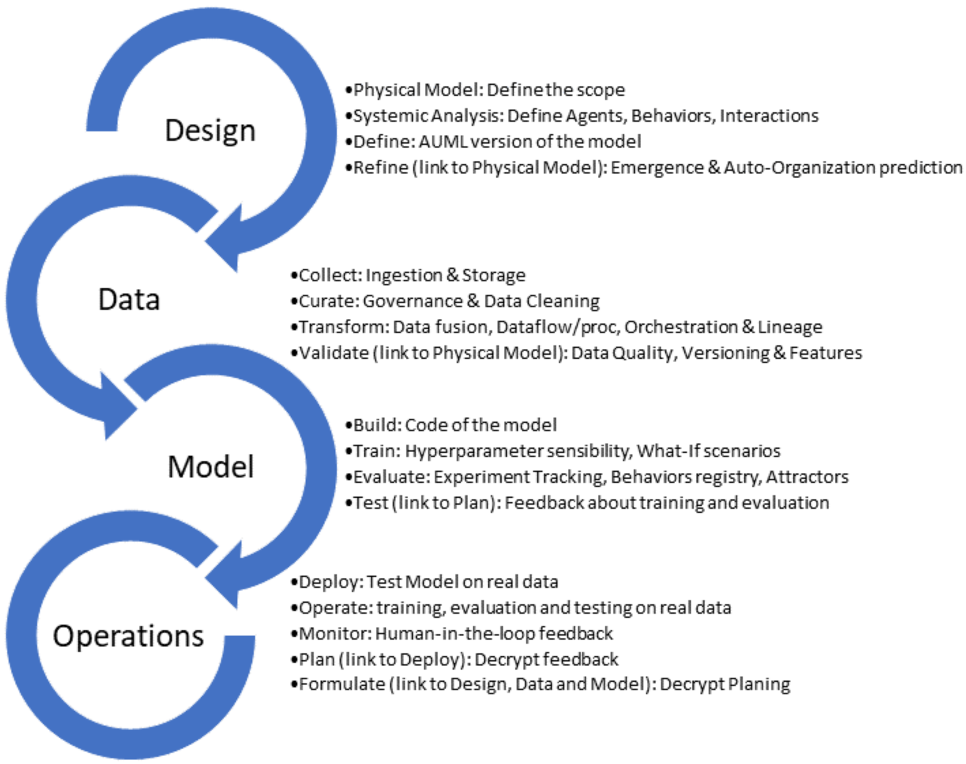

Figure 4: DTOps: Digital Twin Operations Methodology.

(machine monitors, human operators and human supervisors) where each iteration builds upon the previous one, incorporating stakeholders feedback, lessons learned, and evolving requirements. Human-in-the-loop simulation enhances the capabilities of DTs by incorporating human expertise, judgment, and decision-making into the simulation process. It allows decision-makers and operators to interact actively with the DT, observe its behavior, and make informed decisions based on the simulation outcomes.

This interactive loop between the human operator and the DT creates a feedback mechanism that can lead to improved system performance, increased situational awareness, and enhanced decision-making. Human-in-the-loop simulation offers use cases for the DT: decision support, skill development, risk mitigation, stakeholder engagement, and continuous improvement as described in the previous section. However, integrating humans into the simulation process also poses challenges. It requires an accurate representation of human decision-making processes, considering cognitive abilities, biases, and subjective factors.

Another essential point of the DTOps is security. Since this paper focuses on the DTOps methodology and the Data knot, we won't provide an extensive explanation of security measures. For an overview of security threats in DT, please refer to Alcaraz and Lopez (2022), and for open challenges related to security in DT, consult Holmes et al. (2021). The lifecycle must integrate the following points at each step and feedback. Security by Design promotes incorporating security controls and best practices into the design and architecture. This ensures that security measures are inherently built into the twin's structure rather than added as an afterthought. It emphasizes the use of automated security testing tools and techniques to continuously scan and analyze the DT's components, identifying potential security weaknesses or issues. DTs often consist of various interconnected components, such as sensors, data sources, analytics engines, and visualization tools. Security must help to maintain a consistent and secure state of these components, enabling effective management of their configurations, updates, and security profiles. Security managers implement strong access control mechanisms and privilege management strategies, ensuring that only authorized individuals or systems can access and modify the DT's design, data, and related resources (dockers/kubernetes, microservices, etc.). It encourages the use of centralized logging and real-time monitoring to enable proactive security management.

The remainder of the paper will focus on the Data knot, specifically addressing ontology and data management challenges. The discussion of the other knot will be reserved for future papers.

4. Ontology and Data Management in DTOps

As declared by Lenzerini (2011), ontologies and data management are interconnected and play complementary roles in managing and organizing data within domains like SGs, smart cities, or smart buildings. The most known SG Architecture Model is made by CEN-CENELEC-ETSI et al. (2012); see Prieto González et al. (2021) for a recent literature review about ontologies in SGs.

Ontologies provide a standardized representation of concepts, relationships, and properties, facilitating semantic interoperability and serving as a foundation for data modeling and representation. They ensure consistency and compatibility when integrating and exchanging data between heterogeneous systems, enabling data integration and interoperability. Ontologies also help ensure data quality and consistency by defining allowed values, data types, and constraints, and they enhance querying and retrieval capabilities by providing a structured and semantically rich representation of the domain.

In other words, ontologies provide a structured framework for capturing domain-specific information and facilitating semantic interoperability. One of the primary benefits of utilizing ontologies in SG applications is their ability to enhance the capabilities and interoperability of DT systems. Ulivi (2019) proposes a systemic view of ontology, which adopt different levels of observation to describe different systemic levels, especially efficient in complex systems like the SG. The intertwine ontologies facilitate semantic interoperability, enabling data from different sources to be integrated and interpreted coherently on each level of operability. Thus, the main challenge of ontologies is its level of complexity.

Concerning the grid, ontologies capture the hierarchical structure and relationships between various components, such as generators, transmission lines, and substations. Ontologies capture information about different types of renewable energy sources, their geographical locations, and their generation capacities. By integrating this knowledge into DT systems, ontologies enable accurate modeling and simulation of renewable energy generation and its integration into the grid like virtual power plants as proposed by Kott and Kott (2019). The most used ontology's technologies are RDF, OWL, and OWL2.

Since the ontologies are part of the DTOps, they must adopt a feedback process. From the works of Gracia et al. (2010), three kinds of feedback occur. Ontology matching use tools to define shared groundings between the two models; those tools determine the good terminology for the stakeholders and the needs, and a good mapping is useful to find missing or superfluous ontological elements. Semantic reasoning detects inconsistencies between hierarchies (same ontology or between ontologies of different systems). Structure comparison analyzes both models to estimate if they can modify the behaviors of the systemic model. The feedback process is summarized in Figure 5.

On the other hand, data management techniques address various challenges related to data acquisition, integration, preprocessing, storage, analytics, fusion, scalability, and real-time capabilities. These techniques leverage ontologies to ensure data quality, enforce validation rules, perform cleans-

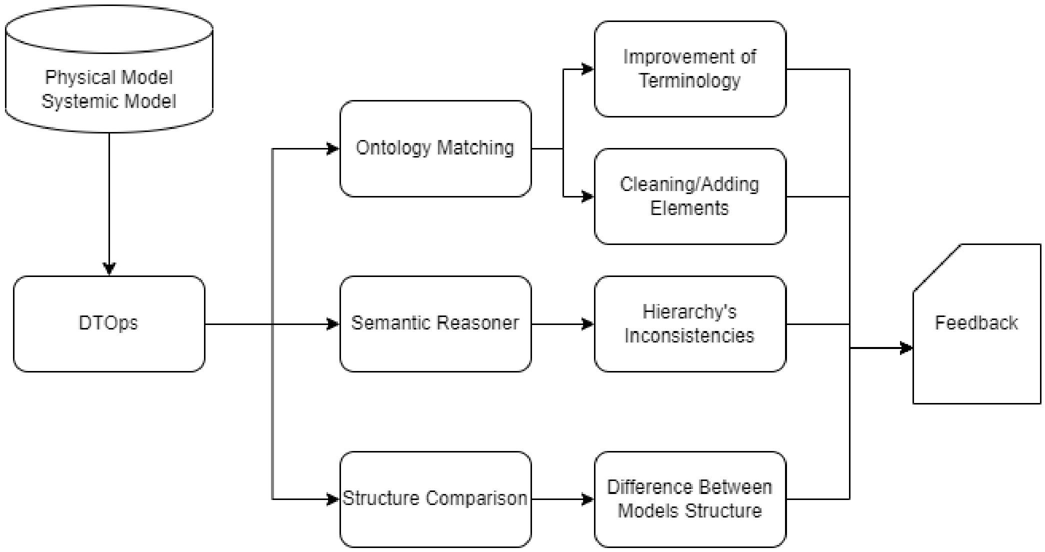

Figure 5: Feedback process for ontologies.

ing and transformation operations, and maintain data integrity and consistency. Data management techniques also utilize ontologies to define query languages, retrieval mechanisms, and analytics algorithms that exploit the semantic relationships captured in the ontology. For a deep understanding of data management characteristics, we refer to Ramakrishnan et al. (2003); Harby and Zulkernine (2022). We provide in Table 1 and Figure 6 a comparison of the main data management and storage technologies and their compatibility with DTs.

With the growing complexity and volume of data generated by SGs, effective data management techniques are essential for accurate representation and analysis within DT environments. Kolajo et al. (2019) describe the main challenges to be related to capturing data from heterogeneous sources, ensuring real-time availability, and handling the large volume and velocity of data generated by SGs.

Singh and El-Kassar (2019) mention that how to integrate the data forms another challenge. The database must ensure data consistency and quality, and leverage ontologies and standardized data models for seamless integration of diverse data sources. Data management must address the scalability requirements, real-time analytics capabilities, and efficient handling of large volumes of data, including distributed storage systems, cloud-based plat-

Aspect	**Data Warehouse**	**Data Lake**	**Data Lakehouse**	**Datamesh**
Concept	Centralized	Centralized	Hybrid	Decentralized
Data Storage	Structured	Structured, Semi-structured, Unstructured	Structured, Semi-struct., Unstructured	Structured, Semi-struct., Unstructured
Data Integration	ETL (Extract, Transform, Load) processes	Schema on read	Schema on read and write	Real-time data streams, Event-driven
Data Processing	Batch processing, SQL-based queries	Flexible, supports batch and real-time processing	Batch processing, real-time processing	Distributed processing, Stream processing
Scalability	Scalable, optimized for structured data	Scalable, handles large volumes of data	Scalable, handles large volumes of data	Scalable, handles Big Data
Flexibility	Limited flexibility, predefined schemas	Flexible, supports diverse data types	Flexible, supports diverse data types	Highly flexible, adaptable to various data sources
Analytics and Insights	Mature analytics capabilities, optimized for structured data	Supports diverse analytics, including structured and unstructured data	Supports diverse analytics, including structured and unstructured data	Enables distributed analytics and insights
Digital Twin Compatibility	Limited compatibility, requires integration	Compatible with digital twin systems, can store and process Digital Twin data	Compatible with digital twin systems, supports integration of those data	Built for seamless integration with digital twin architectures
Use Cases	Business intelligence, reporting	Exploratory analysis, machine learning	Real-time analytics, operational insights	Distributed data mesh, data collaboration

Table 1: Comparison of 4 data management technologies.

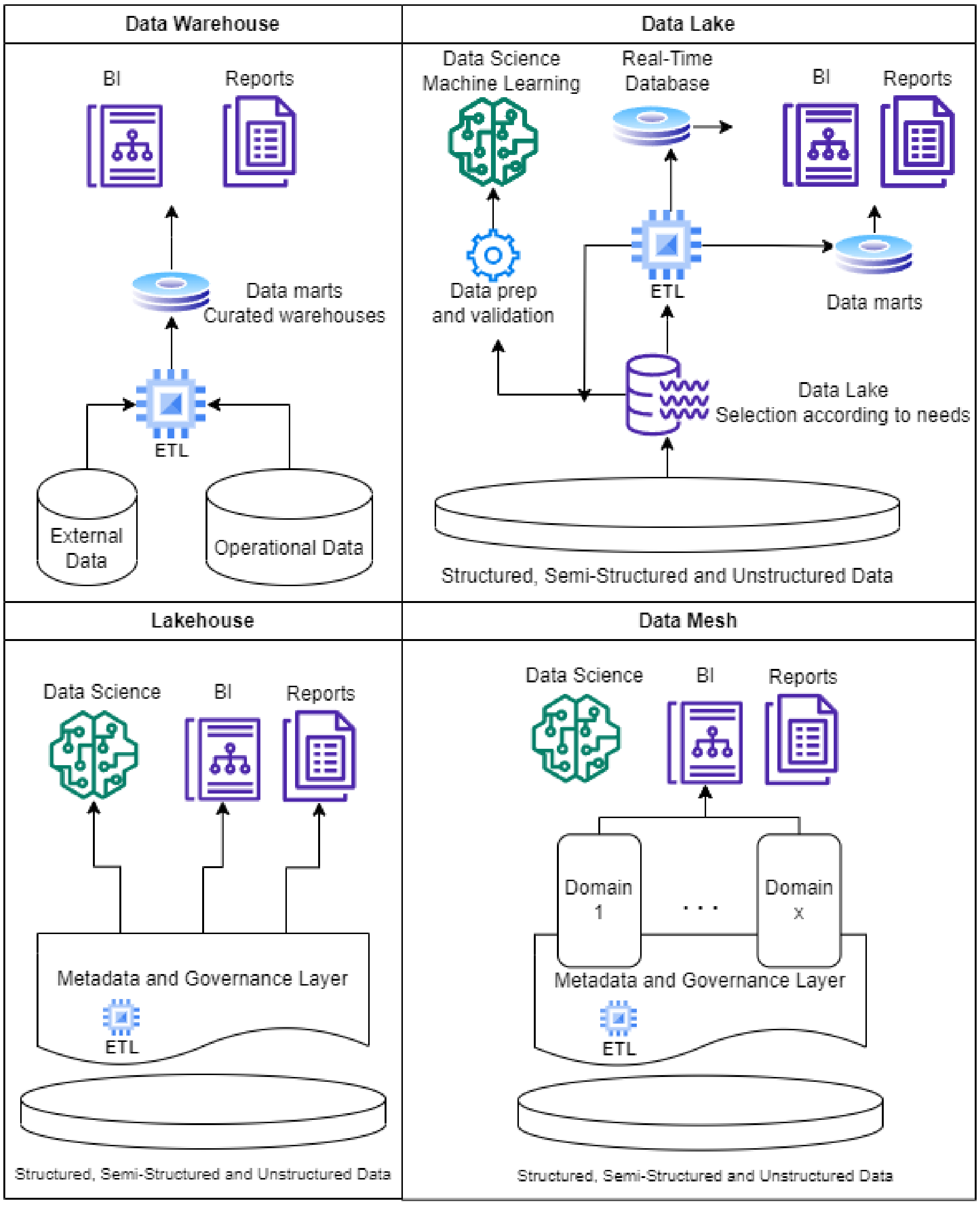

Figure 6: Comparison of 4 data management technologies.

forms, and edge computing technologies. Moreover, the approaches have to handle the growing volume of data, accommodate new devices and sensors, and support the expansion of the DT system without compromising performance or data accessibility.

Once the data is stored, techniques are used to rework the raw data such as data cleaning, filtering, normalization, validation, and anomaly detection to ensure the reliability and accuracy of data used in DT systems (Roh et al. (2019)). Furthermore, stakeholders have to employ techniques such as real-time monitoring, anomaly detection, predictive analytics, and optimization algorithms to derive valuable insights thanks to various technologies like Machine Learning Factory (Lévy et al. (2022)). Another way to work with data is by combining and integrating data from multiple sources to create a more comprehensive and accurate representation of the system (Harby and Zulkernine (2022)), improving situational awareness, fault detection, and predictive capabilities. This process is called the Data Product approach and is the specificities of Data Mesh, Figure 7 illustrates the difference between traditional data management (Top-down) and the data product approach (Bottom-up, or Systemic).

5. Modeling and Applications

Various modeling approaches, such as physics-based models, data-driven models, and hybrid models, are employed to represent the behavior and dynamics of SG systems. Cioara et al. (2021); Sifat et al. (2022); Jafari et al. (2023) list numerous applications of DT in the context of SG from any modeling approaches.

5.1. Definition

Physics-based models utilize fundamental principles and equations to describe the physical processes within the grid. They incorporate mathematical models of power flow, voltage stability, and load forecasting, among others. These models leverage established theories and empirical data to simulate the behavior of the grid components accurately. Physics-based models are valuable for understanding the underlying physics and enabling comprehensive analysis of the SG system.

Data-driven models, on the other hand, rely on historical and real-time data to learn and capture the relationships and patterns within the grid. Machine learning and statistical techniques are applied to analyze large volumes of data and develop models that can predict future system behavior.

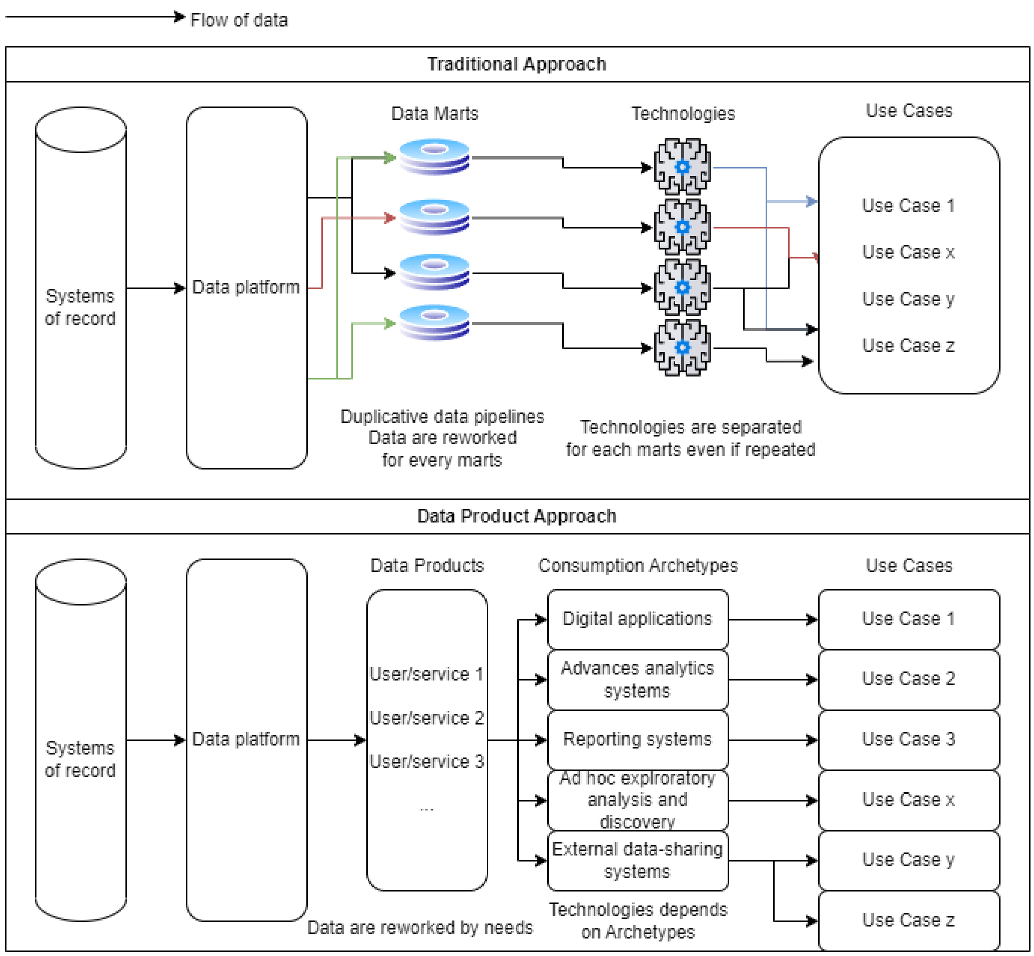

Figure 7: Data Product approach.

Data-driven models are particularly useful for handling complex and non-linear relationships within the grid, enabling accurate forecasting, anomaly detection, and optimization.

Hybrid models, also known as grey-models, combine the strengths of both physics-based and data-driven approaches. They integrate physical equations and empirical data to create comprehensive models that capture the complexities of the SG. Hybrid models offer a balance between accuracy and flexibility, leveraging both domain knowledge and data-driven insights.

A model needs simulations to be valid. Simulation allows for virtual testing of the DT models in various scenarios, simulating different operating conditions and contingencies. It enables the exploration of what-if scenarios and the assessment of system behavior under different parameters and configurations. Simulation provides a cost-effective and safe environment for testing and optimizing SG operations.

Validation is critical to ensure that the DT models accurately represent the physical grid. It involves comparing the behavior and performance of the DT with real-world data and measurements. Validation techniques include comparing simulation results with historical data, conducting field experiments, performing statistical analyses, and including expert validation. Through validation, the reliability and accuracy of the DT models can be assessed, providing confidence in their effectiveness.

By employing modeling and simulation techniques, DT systems for SGs offer invaluable insights into the behavior and performance of the grid. They enable virtual testing, scenario analysis, and performance evaluation, supporting informed decision-making and optimization of grid operations. The combination of physics-based models, data-driven models, and hybrid models empowers DT systems to capture the complexity of the SG accurately. Through simulation and validation, DT models can be refined and improved, ensuring their reliability and usefulness in real-world applications.

We present real-world case studies and examples that showcase the implementation of DT systems in SG applications. These case studies demonstrate the design considerations, ontologies, data management strategies, and modeling and simulation techniques used to develop and deploy DT solutions

5.2. Asset Management and Predictive Maintenance

By creating a virtual replica or representation of physical assets, DTs enable real-time monitoring, analysis, and optimization of asset performance. We refer to the following studies for an in-depth understanding of this concept: Errandonea et al. (2020); Lu et al. (2020b); van Dinter et al. (2022).

DTs provide continuous monitoring of assets, collecting data on their operating conditions, performance, and health parameters. This data can be obtained from sensors, IoT devices, and other monitoring systems. By integrating this real-time data with the DT, operators can gain a comprehensive view of asset behavior and identify potential issues or anomalies. For example, using historical and real-time data, DTs can predict asset degradation and failure patterns. By leveraging advanced analytics and machine learning algorithms, it identifies early warning signs and provides predictive maintenance recommendations.

In other words, DTs enable condition monitoring by comparing real-time asset data with predefined performance thresholds or models. Any deviation from expected behavior can trigger alerts, enabling operators to take timely actions. Condition monitoring helps in identifying potential faults, anomalies, or performance degradation, allowing for proactive maintenance or repair interventions.

By conducting what-if analyses and simulations, operators can evaluate the impact of operational changes or maintenance strategies on asset performance. DTs facilitate the optimization of asset performance by simulating different scenarios, analyzing historical data, and identifying opportunities for improvement. This enables them to make informed decisions to optimize asset utilization, extend asset lifespan, and maximize operational efficiency. It supports the entire lifecycle management of assets. From initial design and installation to operation, asset maintenance, repair, replacement, and retirement. This holistic view enables better asset planning, resource allocation, and decision-making throughout the asset's lifecycle.

5.3. Building Information Modeling

BIM provides a comprehensive and structured digital representation of a building or infrastructure asset, capturing its geometric, spatial, and semantic information. We refer to the following studies for an in-depth understanding of this concept: Lu et al. (2020a); Lee and Lee (2021); Sepasgozar et al. (2023).

The detailed information embedded in BIM models, such as equipment specifications, maintenance schedules, and performance data, can be linked to the DT, allowing real-time monitoring and management of assets. This integration enhances asset visibility, facilitates condition-based maintenance, and optimizes asset utilization as explained in the previous subsection.

BIM models contain detailed information about building components, systems, and energy consumption patterns. By integrating BIM with DTs, energy efficiency analysis can be conducted, to tend to a zero-energy or positive-energy building as proposed by Pereira et al. (2021); Zhao et al. (2021). The DT simulates different energy scenarios, evaluates energy performance, and identifies opportunities for optimization. This supports decision-making for energy-efficient operation and design improvements.

BIM assists in demand response and load management strategies as shown by the works of Esnaola-Gonzalez and Diez (2019). By incorporating real-time data from the SG, the DT simulate and predict energy demand, allowing for proactive load shifting, demand response planning, and optimization of energy distribution. BIM provides the necessary spatial and semantic context to accurately model the building's energy consumption patterns.

BIM models also enable the integration of renewable energy sources and peer-to-peer sharing (Abbasi and Noorzai (2021)). By incorporating solar panels, wind turbines, or other renewable energy technologies into the BIM model, the DT simulates and optimize the energy generation, storage, and distribution within the grid. This integration supports the effective integration of different renewable energy sources and the management of their intermittent nature through batteries and sharing.

The BIM's benefits need to be expanded to a microgrid or a grid.

5.4. Grid Optimization

DTs provide real-time monitoring of energy generation, transmission, and distribution within the SG. By integrating data from sensors, meters, and IoT devices, operators can visualize and analyze energy flows, identify bottlenecks or inefficiencies, and make data-driven decisions to optimize energy usage. DTs enable real-time monitoring and feedback to the physical system. They allow systems to adapt to fluctuations in demand, changes in user specifications, and unexpected disruptions.

DTs enable the modeling and simulation of demand response scenarios; we refer to the review of Teng et al. (2021) for a deeper investigation of this point. By capturing information about consumer profiles, energy demand patterns, and pricing signals, DTs can predict and optimize energy consumption in response to demand fluctuations. This helps balance supply and demand, reduces peak loads, and improves grid stability and reliability. Moreover, operators anticipate future energy needs, optimize energy generation and distribution, and plan maintenance or expansion activities to efficiently meet the growing energy demands. DTs can facilitate parallel controlling of various elements within a system. This means that different parts of a complex system can be optimized simultaneously.

By modeling and simulating the behavior of renewable energy generation, such as solar or wind, DTs help optimize their localization, integration, and utilization (Yu et al. (2022)). This includes forecasting renewable energy production, managing intermittency, and optimizing energy storage and distribution. By creating virtual replicas of microgrid components, including distributed energy resources, energy storage systems, and local loads, DTs enable to optimize energy flows, balance local generation and consumption, and improve the resilience and self-sustainability of microgrids.

DTs allow for simulation and what-if analysis to explore different operational strategies and scenarios. Stakeholders can simulate changes in energy demand, generation capacity, or network configurations to assess their impact on energy optimization. They can assess the energy efficiency of various components, such as transformers, substations, or distribution networks, and identify opportunities for optimization and improvement. This helps in making informed decisions, planning infrastructure upgrades, and optimizing energy resources. Moreover, DTs enable grid planning by simulating different scenarios and evaluating their impact on grid performance. Stakeholders can analyze factors such as load balancing, grid congestion, voltage regulation, and energy losses to optimize grid operations, reduce inefficiencies, and improve overall grid performance.

In summary, DTs rely on extensive data collection and analysis. This data-driven approach provides insights that are not readily available through traditional methods. By continuously analyzing data from sensors and other sources, DTs can identify patterns, inefficiencies, and areas for improvement that might otherwise go unnoticed.

5.5. Planning and Expansion

What-if scenarios are not limited to existing grids or technologies and can also simulate how to change, alter, and enhance the grid. We already mention the use of DTs as feedback to manage or integrate demand-side management and renewable energies; it can also be used for batteries, storage systems, and various infrastructure upgrades. In other words, DTs are useful for grid planning and expansion thanks to the following knowledge:

- Load Forecasting: Coupled with data analysis they are used to predict future electricity demand, enabling utilities to determine the required capacity and infrastructure upgrades.

- DER Integration: Knowledge of DER production is analyzed with various data such as GIS data, population projections, and energy demand forecasts. The optimal placement and capacity of DERs need to be determined to ensure reliable and stable grid operations.

- Equipment obsolescence: Planning and expanding the grid infrastructure involve identifying areas with increased load demand, aging equipment, or potential vulnerabilities thanks to predictive maintenance.
- Policy Framework: Planning and expansion in the SG domain require alignment with regulatory and policy frameworks. Utilities need to consider regulatory requirements, market structures, and incentive programs to support investments in grid expansion. In this case, ontologies play a crucial role.
- Stakeholder Engagement: Collaboration and feedback from various stakeholders help ensure that planning efforts align with the needs and priorities of all involved parties.

Planning for system stability and resilience involves considering factors such as fault tolerance, grid balancing, and contingency planning. Grid expansion strategies should ensure robustness against disturbances and facilitate quick recovery from disruptions, minimizing downtime and improving reliability.

SG planning is a long-term process that involves considering future energy demands, technological advancements, and environmental goals. Utilities need to develop comprehensive roadmaps and investment strategies that account for evolving energy needs and emerging trends.

5.6. Recent work on H2020 MAESHA project

The H2020 MAESHA project focuses on the decarbonization of energy systems on geographical islands. Its primary goal is to facilitate the widespread adoption of renewable energy sources by implementing tailored innovative flexibility services. This is achieved through a comprehensive study and modeling of local energy systems and community structures. MAESHA will initially demonstrate these solutions on the French overseas island of Mayotte and evaluate their potential for replication on five follower islands, representing a combined population of more than 1.2 million inhabitants across both geographical Europe and overseas territories. The project comprises 12 work packages, each described in detail below.

WP1. Study case and requirements, system architecture; It mainly includes the definition of the use-cases, the related requirements, the definition of specific relevant KPIs, the design of the architecture, the definition of the data handling and processing approach (needs, collection, and consolidation), and the set-up of an interoperability-by-design framework.

WP2. Modelling of energy systems and performance forecasting; It led to the creation of an island-scale economy-energy-environment modeling software to be used by partners and local authorities to explore low-carbon medium and long-term energy transition strategies.

WP3. User-centred approach for Local Energy Communities; The social, cultural and economic conditions on Mayotte will be studied and a technologically, economically and socially optimal energy system topology as well as decarbonisation pathway for the island will be developed.

WP4. Energy markets for geographical islands and associated tailored business models; This work package will ensure the commercial viability of the project and determine the business models and costs implication of the developed solutions by setting up an underlying market design and business models for the different market players, aligning the solutions with the local regulatory framework and providing policy and regulatory recommendations for an efficient market uptake in islands.

WP5. Energy Management Systems to enhance the grid flexibility; The general objective of this WP is to design flexibility services that can be offered according to the market design of WP4, specifically tailored towards the reality of geographical islands.

WP6. Additional flexibility through networks synergies improvement and storage; This WP aims at developing solutions able to provide additional flexibility to the grid thanks to networks' synergies improvement and storage systems.

WP7. Communication and control Platform development; The overall objective of this work package is to deploy a utility-scale Control Platform, that will enable efficient management of various flexibilities aggregated by different systems throughout the island.

WP8. Systems integration and validation; The main objective of WP8 is to supervise the integration across all solutions and innovative prototypes.

WP9. Demonstration on Mayotte.

WP10. Replicability study for follower islands and expansion to more islands.

WP11. Communication, dissemination & exploitation of the results.

WP12. Project Management.

In the context of the DTOps design, the MAESHA project's WPs are mapped as shown in Figure 8. While the project was initially defined without the use of a DT, its role remains crucial. WP1, WP2, WP3, and WP4 establish the design requirements. The output of the DT, particularly in the planning and formulation phases, provides valuable insights into how the technologies in WP5, WP6, and WP7 can be integrated into the island without the need for real-world tests. Additionally, given the development of multiple technologies, algorithms, and incentives simultaneously, it becomes challenging to anticipate how the grid will respond without simulation. Once the DT yields the desired behaviors, these technologies can be tested in real conditions. WP8, WP11, and WP12 oversee the development of the DT and facilitate communication among stakeholders across each WP.

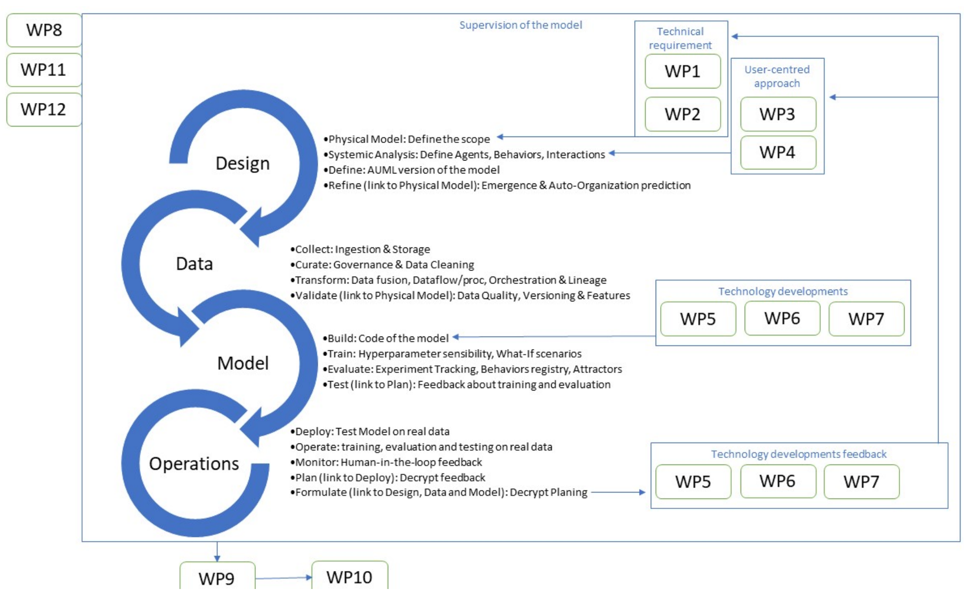

Figure 8: DTOps in MAESHA H2020 project.

6. Bottlenecks

DT technologies hold great promise for SG applications, but they also come with several challenges and limitations that need to be addressed. This section discusses these challenges and highlights potential research directions and opportunities for further advancement in DT design, ontologies, data management, modeling, and simulation in the context of SGs.

Data Management. SGs generate vast amounts of data from diverse sources such as sensors, meters, and devices. Managing and processing this data through ontologies in real-time poses a significant challenge. Ensuring data quality, integrity, security, and scalability are crucial aspects that need to be addressed. Moreover, DTs for SGs are expected to provide real-time monitoring and control capabilities. Ensuring timely data acquisition, processing, and analysis to support real-time decision-making and control actions requires robust infrastructure and advanced algorithms.

Interoperability and Standardization. SGs consist of multiple heterogeneous systems, devices, and applications from various vendors. Ensuring seamless interoperability and integration between these different components can be challenging. Developing standardized interfaces and protocols to enable effective communication and data exchange is essential.

Complexity of Modeling. SGs are complex systems with numerous interconnected components and dependencies. Creating an accurate DT that models and represents the entire SG system can be challenging due to its complexity. Handling the scale of data and ensuring efficient processing and analysis also pose significant challenges. Developing accurate models and validating them against real-world data can be challenging due to the uncertainties, dynamic nature, and diverse operating conditions of the SG. DT models also require regular calibration and updating to maintain accuracy and reliability. As SGs expand and evolve, the DT infrastructure must be scalable to accommodate the increasing number of devices, sensors, and data sources. It should also be flexible enough to adapt to changes in the grid configuration, new technologies, and evolving business requirements.

Definition of the scenarios. With the increasing integration of renewable energy sources and the transition towards a more decentralized grid, DTs must be able to handle the complexities of future grid scenarios. Incorporating advanced modeling techniques, forecasting methods, and optimization algorithms to support the energy transition pose additional challenges.

Collaboration. Successful implementation of DTs requires collaboration and engagement with various stakeholders, including utilities, regulators, technology providers, and end-users. Ensuring effective communication, consensus building, and alignment of objectives among stakeholders can be a complex challenge.

Data Governance and Ethics. DTs rely on extensive data collection and analysis, raising concerns about data governance and ethics. Ensuring transparent and responsible data usage, protecting consumer privacy, and addressing ethical considerations associated with data-driven decision-making are important challenges.

Security and Privacy. Securing the DT infrastructure, communication networks, and data exchange protocols is a significant challenge. Ensuring compliance with privacy regulations and protecting sensitive customer information is also crucial.

Policy. The implementation of DTs may require alignment with existing regulatory and policy frameworks and, on another hand, the successful adoption of DTs relies on the readiness and acceptance of the human workforce. Overcoming resistance to change, providing adequate training, and fostering a culture of data-driven decision-making are challenges that need to be addressed. Adapting regulations to accommodate DT technologies, addressing liability concerns, and ensuring compliance with data protection and cybersecurity regulations pose challenges in the regulatory domain.

7. Conclusion

In this article, we have explored the role and significance of effective DT design, ontologies, data management, modeling, and simulation in enabling SG applications. Through our analysis, we have uncovered key findings and gained valuable insights into the potential of DT technologies for the future development and optimization of SG systems.

First and foremost, we have recognized that DT technologies offer a transformative approach to understanding and managing SGs. By creating virtual replicas of physical systems, DTs enable real-time monitoring, modeling, and control, leading to enhanced operational efficiency, reliability, and sustainability.

The design considerations for DT systems in SGs have emerged as crucial factors for success. We must carefully consider system architecture, data acquisition, integration of heterogeneous data sources, and scalability to develop effective DT systems that accurately represent the complexities of the SG.

Ontologies play a vital role in organizing knowledge and facilitating semantic interoperability in DT systems. By leveraging existing ontologies specific to SG domains, we can enhance the capabilities of DTs and enable seamless integration with various applications and data sources, ultimately improving decision-making processes and system performance.

From data acquisition to preprocessing, fusion, storage, and analytics, managing vast amounts of SG data allows for valuable insights and informed decision-making. Advanced data analytics techniques, such as real-time analytics, machine learning, and predictive analytics, further enhance the value derived from DT systems.

Accurate system modeling enables a deep understanding of the SG behavior, while simulation facilitates virtual testing, scenario analysis, and performance evaluation. The validation of DT models ensures their reliability and usefulness in supporting decision-making processes.

In conclusion, we provide an effective design of DT systems, called DTOps, utilization of ontologies, efficient data management practices, and accurate modeling and simulation techniques.

Aknowledgement

Thanks to the May 24 Version of ChatGPT, this article has undergone enhancements in terms of English coherence and language proficiency.

Fundings

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 957843.

8. CRediT

Conceptualization: GG; Investigation: GG, IT; Methodology: GG, IT; Supervision: GG, SD; Validation: GG, SD; Roles/Writing - original draft:

GG; Writing - review & editing: GG.

References

Abbasi, S., Noorzai, E., 2021. The bim-based multi-optimization approach in order to determine the trade-off between embodied and operation energy focused on renewable energy use. Journal of Cleaner Production 281, 125359.

Ahat, M., Amor, S.B., Bui, M., Bui, A., Guérard, G., Petermann, C., 2013. Smart grid and optimization .

Akbar, M.A., Smolander, K., Mahmood, S., Alsanad, A., 2022. Toward successful devsecops in software development organizations: A decision-making framework. Information and Software Technology 147, 106894.

Alcaraz, C., Lopez, J., 2022. Digital twin: A comprehensive survey of security threats. IEEE Communications Surveys & Tutorials 24, 1475–1503.

Allam, Z., Dhunny, Z.A., 2019. On big data, artificial intelligence and smart cities. Cities 89, 80–91.

Amin, S.M., 2011. Smart grid: Overview, issues and opportunities. advances and challenges in sensing, modeling, simulation, optimization and control. European Journal of Control 17, 547–567.

Amin, S.M., Wollenberg, B.F., 2005. Toward a smart grid: power delivery for the 21st century. IEEE power and energy magazine 3, 34–41.

Amor, S.B., Bui, A., Guérard, G., 2014. A context-free smart grid model using complex system approach, in: 2014 IEEE/ACM 18th International Symposium on Distributed Simulation and Real Time Applications, IEEE. pp. 147–154.

Amor, S.B., Guerard, G., Levy, L.N., 2019. Systemic approach for modeling a generic smart grid, in: Proceedings of the 10th International Symposium on Information and Communication Technology, pp. 15–22.

Attaran, M., Celik, B.G., 2023. Digital twin: Benefits, use cases, challenges, and opportunities. Decision Analytics Journal , 100165.

Bhatti, G., Mohan, H., Singh, R.R., 2021. Towards the future of smart electric vehicles: Digital twin technology. Renewable and Sustainable Energy Reviews 141, 110801.

CEN-CENELEC-ETSI, S.G.C., et al., 2012. Smart grid reference architecture.

Cioara, T., Anghel, I., Antal, M., Salomie, I., Antal, C., Ioan, A.G., 2021. An overview of digital twins application domains in smart energy grid. arXiv preprint arXiv:2104.07904 .

Daki, H., El Hannani, A., Aqqal, A., Haidine, A., Dahbi, A., 2017. Big data management in smart grid: concepts, requirements and implementation. Journal of Big Data 4, 1–19.

De Nicola, A., Villani, M.L., 2021. Smart city ontologies and their applications: a systematic literature review. Sustainability 13, 5578.

van Dinter, R., Tekinerdogan, B., Catal, C., 2022. Predictive maintenance using digital twins: A systematic literature review. Information and Software Technology , 107008.

Dooley, K.J., 1997. A complex adaptive systems model of organization change. Nonlinear dynamics, psychology, and life sciences 1, 69–97.

Egert, R., Grube, T., Volk, F., Mühlhäuser, M., 2021. Holonic system model for resilient energy grid operation. Energies 14, 4120.

Errandonea, I., Beltrán, S., Arrizabalaga, S., 2020. Digital twin for maintenance: A literature review. Computers in Industry 123, 103316.

Esnaola-Gonzalez, I., Diez, F.J., 2019. Integrating building and iot data in demand response solutions, in: Proceedings of the 7th Linked Data in Architecture and Construction Workshop (LDAC 2019), CEUR. pp. 92–105.

Farsi, M., Daneshkhah, A., Hosseinian-Far, A., Jahankhani, H., et al., 2020. Digital twin technologies and smart cities. Springer.

Francisco, A., Mohammadi, N., Taylor, J.E., 2020. Smart city digital twin–enabled energy management: Toward real-time urban building energy benchmarking. Journal of Management in Engineering 36, 04019045.

Freitas, A., Curry, E., 2016. Big data curation. New horizons for a data-driven economy: A roadmap for usage and exploitation of big data in Europe , 87–118.

Fuller, A., Fan, Z., Day, C., Barlow, C., 2020. Digital twin: Enabling technologies, challenges and open research. IEEE access 8, 108952–108971.

Gracia, J., Liem, J., Lozano, E., Corcho, O., Trna, M., Gómez-Pérez, A., Bredeweg, B., 2010. Semantic techniques for enabling knowledge reuse in conceptual modelling, in: The Semantic Web–ISWC 2010: 9th International Semantic Web Conference, ISWC 2010, Shanghai, China, November 7-11, 2010, Revised Selected Papers, Part II 9, Springer. pp. 82–97.

Guérard, G., Ben Amor, S., Bui, A., 2012. A complex system approach for smart grid analysis and modeling, in: Advances in Knowledge-Based and Intelligent Information and Engineering Systems. IOS Press, pp. 788–797.

Guerard, G., Pousseur, H., 2020. Jade modeling for generic microgrids, in: Agents and Multi-agent Systems: Technologies and Applications 2019: 13th KES International Conference, KES-AMSTA-2019 St. Julians, Malta, June 2019 Proceedings, Springer. pp. 377–386.

Harby, A.A., Zulkernine, F., 2022. From data warehouse to lakehouse: A comparative review, in: 2022 IEEE International Conference on Big Data (Big Data), IEEE. pp. 389–395.

Holmes, D., Papathanasaki, M., Maglaras, L., Ferrag, M.A., Nepal, S., Janicke, H., 2021. Digital twins and cyber security–solution or challenge?, in: 2021 6th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM), IEEE. pp. 1–8.

Hribernik, K., Cabri, G., Mandreoli, F., Mentzas, G., 2021. Autonomous, context-aware, adaptive digital twins—state of the art and roadmap. Computers in Industry 133, 103508.

Jacoby, M., Usländer, T., 2020. Digital twin and internet of things—current standards landscape. Applied Sciences 10, 6519.

Jafari, M., Kavousi-Fard, A., Chen, T., Karimi, M., 2023. A review on digital twin technology in smart grid, transportation system and smart city: Challenges and future. IEEE Access .

Jones, D., Snider, C., Nassehi, A., Yon, J., Hicks, B., 2020. Characterising the digital twin: A systematic literature review. CIRP Journal of Manufacturing Science and Technology 29, 36–52.

Kandaperumal, G., Srivastava, A.K., 2020. Resilience of the electric distribution systems: concepts, classification, assessment, challenges, and research needs. IET Smart Grid 3, 133–143.

Khajavi, S.H., Motlagh, N.H., Jaribion, A., Werner, L.C., Holmström, J., 2019. Digital twin: vision, benefits, boundaries, and creation for buildings. IEEE access 7, 147406–147419.

Kolajo, T., Daramola, O., Adebiyi, A., 2019. Big data stream analysis: a systematic literature review. Journal of Big Data 6, 47.

Kott, J., Kott, M., 2019. Generic ontology of energy consumption households. Energies 12, 3712.

Kreuzberger, D., Kühl, N., Hirschl, S., 2023. Machine learning operations (mlops): Overview, definition, and architecture. IEEE Access .

Lee, D., Lee, S., 2021. Digital twin for supply chain coordination in modular construction. Applied Sciences 11, 5909.

Leite, L., Rocha, C., Kon, F., Milojicic, D., Meirelles, P., 2019. A survey of devops concepts and challenges. ACM Computing Surveys (CSUR) 52, 1–35.

Lenzerini, M., 2011. Ontology-based data management, in: Proceedings of the 20th ACM international conference on Information and knowledge management, pp. 5–6.

Lévy, L.N., Bosom, J., Guerard, G., Amor, S.B., Bui, M., Tran, H., 2022. Devops model appproach for monitoring smart energy systems. Energies 15, 5516.

Lodetti, P.Z., João, D.V., Martins, M.A.I., Almeida, J.F., 2021. Remote assistance through augmented reality for the electric power workforce, in: 2021 Asian Conference on Innovation in Technology (ASIANCON), IEEE. pp. 1–6.

Lu, Q., Xie, X., Heaton, J., Parlikad, A.K., Schooling, J., 2020a. From bim towards digital twin: Strategy and future development for smart asset management. Service Oriented, Holonic and Multi-agent Manufacturing Systems for Industry of the Future: Proceedings of SOHOMA 2019 9 , 392–404.

Lu, Q., Xie, X., Parlikad, A.K., Schooling, J.M., 2020b. Digital twin-enabled anomaly detection for built asset monitoring in operation and maintenance. Automation in Construction 118, 103277.

Lv, Z., Chen, D., Lv, H., 2022. Smart city construction and management by digital twins and bim big data in covid-19 scenario. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 18, 1–21.

Lyko, K., Nitzschke, M., Ngonga Ngomo, A.C., 2016. Big data acquisition. New horizons for a data-driven economy: A roadmap for usage and exploitation of big data in Europe , 39–61.

Mbungu, N.T., Bansal, R.C., Naidoo, R.M., Bettayeb, M., Siti, M.W., Bipath, M., 2020. A dynamic energy management system using smart metering. Applied Energy 280, 115990.

Mendi, A.F., 2022. A digital twin case study on automotive production line. Sensors 22, 6963.

Min, Q., Lu, Y., Liu, Z., Su, C., Wang, B., 2019. Machine learning based digital twin framework for production optimization in petrochemical industry. International Journal of Information Management 49, 502–519.

Nguyen, V.K., Sheng, Q.Z., Mahmood, A., Zhang, W.E., Vo, T.D., 2020. Helibot–a smart distributed energy resources platform for futuristic smart grids, in: 2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID), IEEE. pp. 898–901.

Niazi, M.A., Hussain, A., 2012. Cognitive agent-based computing-I: a unified framework for modeling complex adaptive systems using agent-based & complex network-based methods. Springer Science & Business Media.

Olszewski, R., Cegielka, M., Wesolowski, J., 2019. The concept and development of a serious game „alter eco" as part of creating a digital twin of a smart city, in: Entertainment Computing and Serious Games: First IFIP TC 14 Joint International Conference, ICEC-JCSG 2019, Arequipa, Peru, November 11–15, 2019, Proceedings 1, Springer. pp. 426–430.

Pereira, V., Santos, J., Leite, F., Escórcio, P., 2021. Using bim to improve building energy efficiency–a scientometric and systematic review. Energy and Buildings 250, 111292.

Prieto González, L., Fensel, A., Gómez Berbís, J.M., Popa, A., de Amescua Seco, A., 2021. A survey on energy efficiency in smart homes and smart grids. Energies 14, 7273.

Psarommatis, F., May, G., 2023. A literature review and design methodology for digital twins in the era of zero defect manufacturing. International Journal of Production Research 61, 5723–5743.

Ramakrishnan, R., Gehrke, J., Gehrke, J., 2003. Database management systems. volume 3. McGraw-Hill New York.

Ramu, S.P., Boopalan, P., Pham, Q.V., Maddikunta, P.K.R., Huynh-The, T., Alazab, M., Nguyen, T.T., Gadekallu, T.R., 2022. Federated learning enabled digital twins for smart cities: Concepts, recent advances, and future directions. Sustainable Cities and Society 79, 103663.

Reis, J., Housley, M., 2022. Fundamentals of Data Engineering. " O'Reilly Media, Inc.".

Roh, Y., Heo, G., Whang, S.E., 2019. A survey on data collection for machine learning: a big data-ai integration perspective. IEEE Transactions on Knowledge and Data Engineering 33, 1328–1347.

Rothrock, L., Narayanan, S., 2011. Human-in-the-loop simulations. Springer.

Sakhnini, J., Karimipour, H., Dehghantanha, A., Parizi, R.M., Srivastava, G., 2021. Security aspects of internet of things aided smart grids: A bibliometric survey. Internet of things 14, 100111.

Schroeder, G.N., Steinmetz, C., Rodrigues, R.N., Henriques, R.V.B., Rettberg, A., Pereira, C.E., 2020. A methodology for digital twin modeling and deployment for industry 4.0. Proceedings of the IEEE 109, 556–567.

Sepasgozar, S.M., Khan, A.A., Smith, K., Romero, J.G., Shen, X., Shirowzhan, S., Li, H., Tahmasebinia, F., 2023. Bim and digital twin for developing convergence technologies as future of digital construction. Buildings 13, 441.

Sifat, M.M.H., Choudhury, S.M., Das, S.K., Ahamed, M.H., Muyeen, S., Hasan, M.M., Ali, M.F., Tasneem, Z., Islam, M.M., Islam, M.R., et al., 2022. Towards electric digital twin grid: Technology and framework review. Energy and AI , 100213.

Singh, M., Fuenmayor, E., Hinchy, E.P., Qiao, Y., Murray, N., Devine, D., 2021a. Digital twin: Origin to future. Applied System Innovation 4, 36.

Singh, S., Shehab, E., Higgins, N., Fowler, K., Reynolds, D., Erkoyuncu, J.A., Gadd, P., 2021b. Data management for developing digital twin ontology model. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 235, 2323–2337.

Singh, S.K., El-Kassar, A.N., 2019. Role of big data analytics in developing sustainable capabilities. Journal of cleaner production 213, 1264–1273.

Studer, S., Bui, T.B., Drescher, C., Hanuschkin, A., Winkler, L., Peters, S., Müller, K.R., 2021. Towards crisp-ml(q): A machine learning process model with quality assurance methodology. Machine Learning and Knowledge Extraction 3, 392–413. URL: https://www.mdpi.com/2504-4990/3/2/20, doi:10.3390/make3020020.

Taleb., I., Guerard., G., Fauberteau., F., Nguyen., N., 2023. A holonic multi-agent architecture for smart grids, in: Proceedings of the 15th International Conference on Agents and Artificial Intelligence - Volume 1: ICAART, INSTICC. SciTePress. pp. 126–134. doi:10.5220/0011803300003393.

Tao, F., Qi, Q., Wang, L., Nee, A., 2019. Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: Correlation and comparison. Engineering 5, 653–661.

Tao, F., Xiao, B., Qi, Q., Cheng, J., Ji, P., 2022. Digital twin modeling. Journal of Manufacturing Systems 64, 372–389.

Teng, S.Y., Touš, M., Leong, W.D., How, B.S., Lam, H.L., Máša, V., 2021. Recent advances on industrial data-driven energy savings: Digital twins and infrastructures. Renewable and Sustainable Energy Reviews 135, 110208.

Ulivi, L.U., 2019. First steps toward a systemic ontology. Systemics of Incompleteness and Quasi-Systems , 57–73.

Wang, B., Zhou, H., Yang, G., Li, X., Yang, H., 2022. Human digital twin (hdt) driven human-cyber-physical systems: Key technologies and applications. Chinese Journal of Mechanical Engineering 35, 11.

Yitmen, I., Alizadehsalehi, S., Akıner, İ., Akıner, M.E., 2021. An adapted model of cognitive digital twins for building lifecycle management. Applied Sciences 11, 4276.

Yu, W., Patros, P., Young, B., Klinac, E., Walmsley, T.G., 2022. Energy digital twin technology for industrial energy management: Classification, challenges and future. Renewable and Sustainable Energy Reviews 161, 112407.

Zagan, E., Danubianu, M., 2020. Data lake approaches: A survey, in: 2020 International Conference on Development and Application Systems (DAS), IEEE. pp. 189–193.

Zhao, L., Zhang, H., Wang, Q., Wang, H., 2021. Digital-twin-based evaluation of nearly zero-energy building for existing buildings based on scan-to-bim. Advances in Civil Engineering 2021, 1–11.

Zheng, X., Lu, J., Kiritsis, D., 2022. The emergence of cognitive digital twin: vision, challenges and opportunities. International Journal of Production Research 60, 7610–7632.