

Digital Twins: A Meta-Review on Their Conceptualization, Application, and Reference Architecture

Alexander Rossmann
Reutlingen University
alexander.rossmann@reutlingen-university.de

Dieter Hertweck
Reutlingen University
dieter.hertweck@reutlingen-university.de

Abstract

The concept of digital twins (DTs) is receiving increasing attention in research and management practice. However, various facets around the concept are blurry, including conceptualization, application areas, and reference architectures for DTs. A review of preliminary results regarding the emerging research output on DTs is required to promote further research and implementation in organizations. To do so, this paper asks four research questions: (1) How is the concept of DTs defined? (2) Which application areas are relevant for the implementation of DTs? (3) How is a reference architecture for DTs conceptualized? and (4) Which directions are relevant for further research on DTs? With regard to research methods, we conduct a meta-review of 14 systematic literature reviews on DTs. The results yield important insights for the current state of conceptualization, application areas, reference architecture, and future research directions on DTs.

1. Introduction

The concept of digital twins (DTs) is receiving increasing attention in research and management practice [1], [2]. Mentioned initially in 2003 by Michael Grieves, DTs pave the way for cyber-physical integration and serve as a bridge between the physical world and the cyber world [3]. Grieves's initial description defines a DT as a virtual representation of a physical product [4]. Later, Grieves expands on this definition by describing DTs as consisting of three components: a physical product, a virtual representation of that product, and bidirectional data connections that feed data from the physical product to the virtual representation and back [4]. Scholars have also provided evidence for the notion that DTs not only deal with physical products but also are applicable to any

connection and bidirectional exchange between a virtual and physical entity [5]–[7].

Initially conceptualized for manufacturing [3], [8], the idea of DTs has reached multiple domains, including smart cities [9], healthcare [10], management [11], and psychology [12]. Since the inception of DTs in 2003, the concept has grown in interest and was listed as a key strategic technology trend for the first time by Gartner in 2019.¹ A search of the term digital twin in academic databases shows increased research interest in the topic. Until 2017, the number of academic articles on DTs was only in the single-digit range. Since then, the number of yearly publications has grown exponentially [13].

The idea of DTs defined by cyber-physical integration is creating a broad array of opportunities for organizations. However, owing to the short and dynamic development cycle, the concept of DTs in terms of a clear conceptualization of the concept and its properties, an overview of application areas and enabling technologies, and reference architectures for the implementation of DTs is still blurry [1], [2]. In response to the ambiguity in current research, this paper aims to answer four research questions: (1) How is the concept of DTs defined? (2) Which application areas are relevant for the implementation of DTs? (3) How is a reference architecture for DTs conceptualized? and (4) Which directions are relevant for further research on DTs?

The approach for the exploration of these research questions follows established guidelines for a meta-analysis on systematic literature reviews (SLRs) [14], [15]. Therefore, we conducted a meta-review of 14 existing SLRs on DTs and analyzed the body of related work with respect to the presented research scope. Accordingly, we present and discuss the results of the review. The results yield important insights for the current state of conceptualization, application areas, reference architecture, and further research on DTs.

¹ Gartner Hype Cycle for Emerging Technologies 2019, <https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2019/>.

2. Related work

A search on Google Scholar and Web of Science with the search strings (Systematic Literature Review) AND (Digital Twin*) on the topic level leads to 14 articles, with related work published between 2018 and 2021. The following review is fueled by a full read of these articles and a corresponding analysis of the areas of application, time frames, databases, and article quantity of each SLR, as well as an overview of central concepts of each article.

Multiple reviews on the topic of DTs are present in current research. Initially, the DT concept emerged in the context of *manufacturing* and *Industry 4.0* [16]. Kritzinger et al. [17] presented an SLR on DTs in manufacturing in 2018. They noted that DTs are “commonly known as a key enabler for the digital transformation, however, in literature is no common understand concerning this term”. In this review, the authors adopt the definition of Tao et al. [18] that “the digital twin is an integrated multi-physics, multi-scale, probabilistic simulation of a complex product and uses the best available physical models, sensor updates, etc., to mirror the life of its corresponding twin”. An important contribution of Kritzinger et al. [17] is the conceptualization of DTs with respect to the level of integration between the physical and its digital counterpart. The terms “digital model,” “digital shadow,” and “digital twin” are often used synonymously. However, the given definitions differ in the level of integration; that is, a digital model contains a pure digital representation of an existing or planned physical object that does not use any form of automated data exchange, and a digital shadow comprises a one-way data flow from the state of an existing physical object to a digital object. Adopting this distinction, Kritzinger et al. [17] categorized 53 articles published between 2001 and 2017 with respect to the application of digital models, digital shadows, and DTs.

Comparable reviews with a strong focus on manufacturing and Industry 4.0 were conducted by Cinar et al. [16] and Catarci et al. [19] with a rather low number of articles ($n = 19$ and $n = 16$, respectively). A major contribution of Cinar et al. [16] is their work on GE digital with respect to a hierarchical classification of DTs in four properties: component, asset, system, and process. A component twin is a major sub-component affecting the performance of the asset to which it belongs. Asset twins can be a collection of component twins, such as a motor or pump. System twins are a collection of asset twins performing a system function such as a production line. Process twins usually provide a perspective to a set of operations at the highest level and generally focus on processes rather than equipment.

The main contribution of Catarci et al. [19] is the connection of concepts dealing with cyber-physical systems (CPSs), the Internet of Things (IoT), and DTs in smart manufacturing. They envision a system architecture for DTs in which humans can specify a goal and take advantage of DTs to automatically compose the corresponding physical processes. An important aspect of this view is the introduction of a service perspective to DTs. Various properties of DTs are discussed (e.g., connectivity, autonomy, homogeneity, ease of customization, traceability). Moreover, Catarci et al. [19] define data models, patterns for data exchange, and the ability to run simulations as core facets of DTs and review current industry solutions for DTs in smart manufacturing as Eclipse Ditto, Bosch IoT Things solutions, GE Predix, Microsoft Azure IoT, Amazon AWS IoT, and IBM Watson IoT.

A corresponding SLR in the area of smart production is the concept of cyber-physical production systems (CPPSs) [5]. CPPSs are systems of systems with autonomous elements connected with each other, on and across all levels of production, from processes through machines up to production and logistics networks, enhancing decision-making processes in real time, response to unforeseen conditions, and evolution along time. Wu et al. [5] reviewed 100 articles with respect to a conceptual definition and engineering development of the topic. They arrived at a concept map with three conceptual categories of articles: need analysis, concept exploration, and concept definition. Furthermore, they defined five categories of articles on an engineering level: smart connectivity, data-to-information conversion, cyber, cognition, and configuration. In particular, DTs in this framework are part of the cyber level on the engineering side of CPPSs.

Lim et al. [20] provide a comprehensive overview of the technology stack for DTs in smart manufacturing. They also discuss potential contributions of DTs along the whole product life-cycle management process and expand perspectives to business models and business innovation. Based on an SLR of 123 research articles published between 2015 and 2019, the technology stack of DTs covers multiple layers, including data acquisition, data management and connectivity, network architecture, data representation and storage tools, data analytics and machine learning (ML), microservices, and cyber security. Furthermore, Lim et al. [20] discuss contributions of the DT concept on different product life-cycle management stages as product design, manufacturing, distribution, usage, and end of life. This allows various fresh perspectives for business model innovations, covering aspects of corporate strategy, customer and market segmentation, and value creation components.

With the further development of research on DTs, increasingly more scholars are working on an extension of the concept beyond manufacturing. An example of an SLR in this direction is that of Zhang et al. [6]. They reviewed 59 articles on DTs and services to develop a framework for holistic industrial product service systems. Within such systems, the application of DTs is reviewed along the whole life cycle of products and services from beginning to end of life. This includes a notion on design, sales, distribution and usage, and reuse and recycling of products and services.

A comparable review extends the analysis from smart production to industrial services with an additional focus on predictive maintenance and after-sales services. Melesse et al. [21] summarize 25 research articles in the selected domains and argue that DTs play an important role in operations throughout the whole product life cycle, including the concept and design stage, manufacturing planning and execution, sales, product usage, maintenance, and product renewals.

Beyond the area of smart manufacturing, several other SLRs extend the idea of DTs to a more general level of conceptualization. Jones et al. [22] characterized DTs through a review of 92 articles and identified 12 characteristics. Core characteristics of DTs include *physical entities* (e.g., vehicles, components, products, artefacts); *virtual entities* (e.g., data, models); the *physical environment* of a DT (e.g., factories, cities); the *virtual environment* of a DT (e.g., databases, data warehouse, cloud platforms, servers); *parameters*, as types of data, information, and processes that are passed between the physical and virtual entities; *fidelity*, defined as the number of parameters, their accuracy, level of abstraction, and transfer between the virtual and physical environment; *state*, or the current condition of both the physical and virtual twins or the current values for each of the measured parameters; *physical-to-digital* versus *virtual-to-physical connectivity*; *twinning rate*, or the act of synchronization between the virtual and physical states (e.g., real-time, near-time); *physical processes* (e.g., smart factories, three-dimensional (3D) printing, robot control, medical health applications); and *virtual processes* (e.g., simulation, modeling, optimization).

A comparable SLR for general conceptualization conducted by Enders and Hossbach [23] contains an analysis of the dimensions for DT applications. With a review of 152 research articles, they separated *purpose*, *industrial sector*, the *physical reference object* for the DTs, and other relevant dimensions for the description of current applications. By 2019, most applications were linked to manufacturing, while other sectors such as automotive, aerospace, energy, healthcare, and smart cities were also working with DTs. The main purpose of

most applications is simulation, directly followed by monitoring and control. The idea of DTs is applicable to various industries with multiple application areas.

Enders and Hossbach [23] also argued for a deeper exploration of the DT concept in the information systems domain and postulated the need for a corresponding taxonomy. Van der Valk et al. [13] describe such a taxonomy of DTs based on an SLR of 233 articles from different databases. A taxonomy describes properties of a research object and relevant differences due to specification of these properties in research and practice. Relevant properties of DTs include the data link between the physical and virtual parts of the DTs (unidirectional, bidirectional), purpose (processing, transfer, repository), the connection between the physical and virtual parts (physically independent, physically bound), accuracy of data exchange (identical, partial), interface (machine-to-machine, human-to-machine), synchronization, data input (raw, processed), and the time the physical and virtual parts are created (physical part first, virtual part first, simultaneously).

In addition to the focused reviews in smart manufacturing and holistic reviews on taxonomies, properties, and general characteristics of DTs, other SLRs have explored the concept within a specific context or sub-topic. Dos Santos et al. [24] examined the application of simulations for decision support in DTs. In a review of 75 articles, they analyzed different application areas and objectives for decision-making through simulations with DTs. The main application areas for such simulations are manufacturing, services, logistics, healthcare, and constructions. Regarding decision-making objectives, simulations with DTs are used for production planning, process evaluation, process control, and resource allocation. Dos Santos et al. [24] also researched the applied platforms for simulations such as Tecnomatix or Arena and relevant software frameworks such as Python, Java, and Stroboscope. Moreover, they evaluated different types of connectivity between the simulation model and the physical system (e.g., IoT devices, web services, management systems), updating and synchronization practices (real-time, near real-time), and the degree of autonomy of the DT simulation model (e.g., autonomous command, recommended actions).

Another SLR with a more tapered orientation by Rathore et al. [25] captures the role of big data and ML in digital twinning. In a review of 61 sources in various databases, patents, and technical reports, the authors identified different applications of big data and ML in the context of DTs in various industries. Examples of potential ML algorithms and data models include applications in production, healthcare, transportation, education, and business. Furthermore, Rathore et al.

[25] proposed a model for the integration of IoT, big data, ML, and DTs, in which (1) IoT and other data sources create big data, (2) data are employed in data models and ML algorithms, (3) simulations and automation procedures are executed in the virtual environment, and (4) such simulations and automated processes are used for deployments in the physical part of a DT. The SLR offers a detailed analysis of this model with a description of the applied ML approaches in different industries. Furthermore, the authors offer an extensive overview on DT development tools, evaluation procedures for the success of DT applications, and reference architectures for digital twinning.

The concept of DTs is intensively integrated in the development of CPSs, smart production, and the vision of Industry 4.0. However, the basic idea behind DTs is also applicable to other industries. Therefore, further SLRs are available for DTs in smart cities [26] or the enablers of and barriers to DTs in the process industry [27]. In their gray literature analysis, Ketzler et al. [26] compared the concept of DTs with established 3D city simulations. They evaluated commonalities and differences between the two concepts and argued that DTs describe something more than a 3D city model (including semantic data, real-time sensor data, physical models, and simulations). Furthermore, they analyzed current implementations of DTs in cities and concluded that there are significant opportunities for up-scaling DTs, with the potential to bring benefits to the city and its citizens.

Finally, Perno et al. [27] conducted an SLR on enablers of and barriers to the implementation of DTs in the process industry. From a review of 38 articles, they developed a framework for organizational capabilities, several development issues, and performance effects for DTs. As such, organizational preconditions such as knowledge, design, and integration of a DT are prevalent for implementation success. Moreover, the development process itself contains several barriers (e.g., lack of standardization, model development, data quality, IT infrastructure). Finally, performance issues need to be discussed at the very beginning of a DT initiative (e.g., costs vs. benefits, reliability, robustness).

In summary, related SLRs on the concept of DTs have initially appeared in the area of *smart production* and *Industry 4.0*, with the distinction among digital models, digital shadows, and DTs [17]; a hierarchical classification of levels for DTs in a production environment [16]; a conceptual connection among CPSs, IoT, and DTs; and emerging perspectives on system architectures and services based on DTs [19]. Within larger systems of systems, DTs are the digital part of CPPSs [5]. The technology stack and several implications of DTs along the whole product life-cycle

management chain, business models, and business innovation are discussed by Lim et al. [20]. Several scholars have expanded the idea of DTs beyond production. This covers a discussion on the role of DTs in industrial product service systems [6], [21]. Beyond smart production, SLRs tend to develop *general characteristics of DTs* applicable in multiple domains [22]. Therefore, the concept of DTs is applicable to various industries and application areas such as automotive, aerospace, energy, healthcare, and smart cities [23]. A taxonomy can be used to define general properties and common differences in the specification of DTs [13]. Beyond holistic frameworks for DTs are SLRs with a strong focus on a specific area, such as the application of simulations based on DTs [24] or the connection among big data, ML, and DTs [25]. Rathore et al. [25] offer an extensive review on the connection between IoT and other data sources, data models, ML algorithms, and data-based applications on the virtual and physical part of DTs. Such frameworks and generic reference architectures might be transferred to various application areas—for example, to the implementation of DTs for cities [26] or specific applications of DTs in the process industry [27]. Table 1 summarizes the area of application, time frame, database, quantity of research articles, and the general concepts of the different SLR on DTs.

Table 1. SLRs on the concept of DTs

| # | Area of application | Time frame, databases, quantity | Concepts |
|------|---|---|---|
| [17] | Manufacturing, Industry 4.0 | 2001–2017, Google Scholar, Scopus, n=53 | DTs, digital model, digital shadow, enabling technologies |
| [16] | Manufacturing, Industry 4.0 | 2015–2019, database not defined, n=19 | Manufacturing areas, hierarchical classification |
| [19] | Manufacturing, Industry 4.0, digital factory | 2008–2018, Google Scholar, n=16 | CPS, IoT, properties, system architecture |
| [5] | Manufacturing, Industry 4.0, CPPS | 2015–2019, Web of Science, n=100 | CPPS, purpose, concept definition, design, simulation, validation, architecture |
| [28] | Manufacturing, product life-cycle management, business innovation | 2015–2019, Scopus, n=123 | Technology stack, network architecture, data exchange protocols, ontologies, middleware, computational processing, ML, micro-services |

| # | Area of application | Time frame, databases, quantity | Concepts |
|------|---|--|--|
| [6] | Services, industrial product service systems | 2008–2018, Scopus, n=59 | Services, product design, customer purchase and usage, reuse, recycling |
| [21] | Industrial operations, production, predictive maintenance, after-sales services | 2016–2019, Scopus, Web of Science, n=25 | Concept, design, manufacturing planning and execution, sales, product usage, maintenance, product renewals |
| [23] | General, DTs | 2002–2018, Google Scholar, n=152 | Purpose, industrial sector, physical reference object |
| [13] | Taxonomy, properties | na–2020, ACM, AIS, IEEE, JSTOR, Science Direct, n=233 | Taxonomy, data link, purpose, connection, accuracy, interface, synchronization, data input, time of creation |
| [24] | Simulation, decision making, CPS | na–2020, Scopus, Web of Science, Scielo, IEEE, Science Direct, n=75 | Application areas, objectives, platforms, connectivity, autonomy |
| [25] | Artificial intelligence (AI), ML, big data | 2015–2020, IEEE, ACM, Scopus, SpringerLink, Hindawi, IGI, Taylor& Francis, Wiley, n=61 | Standards, industry applications, integrated models of IoT, big data, AI and DTs, evaluation of DTs, development tools, reference architecture |
| [27] | Process industry, enablers, barriers | 2016–2020, Scopus, Web of Science, n=38 | Organizational capabilities, development, integration, performance, security |
| [26] | Smart cities, 3D city models | Gray literature review | 3D city models, application of DTs in cities, challenges, opportunities |

3. Results

We present the results of the meta-review of SLRs in the area of DTs according to the formulated research questions. Therefore, we organize the following subsections around conceptualization, application areas, reference architecture, and future research directions for DTs.

3.1. Conceptualization of DTs

The conceptualization of DTs is an important subject of multiple reviews and coined by the area of application. Early articles on the subject focused on smart manufacturing, physical production, and products [16]. Therefore, such conceptualizations import domain-specific aspects into the definition of a DT. However, given the increasing application of the concept in multiple domains, a general conceptualization without any domain-specific properties is required. Therefore, a mutual understanding of a broad array of reviews defines a DT as a CPS with physical and virtual (digital) parts. As Grieves [4] argued, DTs serve as a bridge between the physical world and the cyber world. Data flows between a physical and a digital object with full integration in both directions can be viewed as a central property of DTs [17]. This feature leads to multiple implications, as twinning demands not just a simple image of the physical object but also a real interaction between the physical and digital parts of the twin. Thus, the conceptualization of DTs needs to be extended by various other properties such as data models, connectivity, accuracy, and synchronization [13], [24]. Moreover, the development of a DT is driven by a specific purpose and expected benefits [6], [23]. DTs are only valuable if the processing of real-world data in the digital part leads to relevant insights and corresponding services in the physical part of the twin. This leads to the assumption that certain applications in the area of AI (e.g., ML algorithms) should be treated as constitutional property of a DT [25].

Given the mentioned properties, DTs need to be conceptualized as a CPS, with an interactive relationship between the physical and digital parts, purpose, data connectivity in both directions, corresponding data models, task-specific levels of model accuracy and data synchronization procedures, embedded AI in the digital part of the twin, and dedicated services in the physical part.

3.2. Areas of application

Such a general conceptualization of DTs is fruitful for multiple application areas. Originally conceptualized for manufacturing [3], [8], over time, the idea of DTs has reached multiple domains, including smart cities [9], healthcare [10], management [11], and psychology [12]. Therefore, the general idea of DTs is not exclusively linked to the specific domain; rather, digital twinning provides a framework for applications in multiple domains and industries with a focus on two-way interactions between a physical and digital entity. Such a generalization of the concept is already

embedded in the analyzed SLRs, in which several authors developed approaches for DTs in smart manufacturing [16], [17], [19], but also in CPPSs with a focus on services in an extended production environment [5], stakeholder-specific services throughout the whole product life cycle [6], or ideas for the application of DTs in marketing and sales [21]. Thus, this meta-review provides evidence for a systematic expansion of the concept in further domains. It is likely that this expansion will continue over the next years with diverse and fruitful applications for DTs in multiple domains.

3.3. Reference architecture

Relevant sub-concepts for a general reference architecture for DTs are present in multiple SLRs within this meta-review [13], [22], [25], [28]. Rathore et al. [25] provide a holistic approach for a general reference architecture for DTs with multiple layers. The different layers constitute a hierarchical order and include multiple forward and backward interactions. Furthermore, a distinction between different layers supports the definition of properties and relevant technologies per layer. Essential components for such an architecture are also presented in Lim et al.'s [20] SLR. Table 2 gives an overview of a nine-layer architecture and corresponding properties. Space restrictions prevent the inclusion of corresponding technologies and application examples for each layer of the architecture in Table 2. However, more precise specifications are available in the assigned SLRs per layer.

The bottom layer for the architecture of a digital twin comprises physical entities [22], which include vehicles, components, products, machines, streets, parking lots, buildings, and other physical objects. The commonality in these entities lies in their real-world existence. To encompass all types, and in line with existing literature [22], we propose the use of the term “physical entity” for general applicability, where a physical entity exists regardless of whether it has been twinned, and the more specific term “physical twin” for when the physical entity is twinned.

The second layer of the architecture includes different strategies for data generation [20]. This requires an identification of the physical entities for digital twinning and the relevant parameters to be generated. Parameters refer to the types of data and the information that can be generated with the data. Typical technologies for data generation include sensors or log-files. In addition, fidelity is a relevant property defined on the data generation layer. The term “fidelity” describes the number of parameters, their accuracy, and the level of abstraction that is transferred between the

virtual and physical twin [22]. The definition of fidelity describes the required accuracy for digital twinning, such as whether a minor part of the physical entity is twinned [29] or the DT is a full mirroring of the physical characteristics and functionalities [30].

The third layer of the reference architecture includes network and connectivity. This layer covers data acquisition and transmission as crucial elements for real-time information flow and connectivity [28]. The layer emphasizes network architectures, data exchange protocols, and middleware platforms to facilitate information exchange and streaming processing. Network architecture involves integration of protocols. Such communication protocols are crucial rule sets for machine-to-machine connectivity between communicating entities. In addition to network architecture, this layer covers the connectivity infrastructure applied in the DTs (e.g., Wi-Fi, Bluetooth, ZigBee, mobile radio communication).

The fourth layer comprises data storage, data integration, and big-data processing. Heterogeneous data sources and domain knowledge gathered from application processes need to be integrated in operational database systems and an integrated data lake [20].

On the fifth layer, the integrated data need to be interpreted and prepared. Knowledge representation tools for DT creation such as ontologies are potential choices for achieving knowledge-based systems. Ontologies are favored because they address integration and domain-specific modeling concerns as well as the reuse and sharing of knowledge [20]. Knowledge representation languages such as the W3C web ontology language (OWL) and knowledge management models such as the resource description framework give the foundation for DT creation, while semantic integration of sensor data is explored to create taxonomies, ontologies, data formats and standards.

Accordingly, the sixth layer is crucial for DTs and deals with data models, algorithms, virtual entities, and virtual twins. ML and data-processing tools provide multiple solutions ranging from analytics to automation, and these provide DTs with decision-aiding capabilities via enabling tools [20].

The seventh layer is dedicated to micro-services and the deployment of data models in real-life processes [20]. Microservices are software development tools constructed as a set of loosely coupled services. This architectural style can be described as an enabling feature for an application to be built as a suite of relatively small, consistent, isolated, and autonomous services performing specific tasks [31]. Microservice architectures are available for different domains; for example, in the application area of smart production, RAMI 4.0 [32] provides an overview of layers and

microservices in production systems to allow monitoring and tracing services of shop-floor assets for automated conflict resolution and performance enhancement through decision-aiding support and control.

The eighth layer deals with system security and data privacy. As DTs affect real-life processes, security is a major and cross-architectural issue within the design of twinning concepts. Owing to the integration of sensitive data in the case of stakeholder-related services (e.g., after-sales, product and content customization), privacy is also becoming a relevant topic. Privacy-preserving approaches in DTs can be classified into two categories: cryptographic approaches featuring encryption schemes and cryptographic primitives and noncryptographic approaches with a policy-based authorization infrastructure [27]. Properties of a reference architecture should be able to detect security and privacy concerns and minimize breaches and associated risks to which DTs can be exposed.

Finally, the ninth layer pertains to the generation of business models. This layer is of great importance because DTs only make sense if they drive business innovations and lead to compelling customer experience and business models [20]. Although interest is growing in multiple application areas, to date, DTs remain predominantly applied in the manufacturing industry. However, even in smart manufacturing, only a few scholars have focused on the business model implications of DTs [33], [34]. DTs strive to enhance customer experience through better-suited products and services. Attaining customer satisfaction through better quality products and services, while enlarging the customer base via new market access and co-creation initiatives, drives the business model for DTs in smart manufacturing. DTs also drive the development of new products, services, and value propositions.

In summary, the twinning process begins with the collection of data from the physical entities or with the usage of already-collected data in the virtual twin (using computer-aided software and/or simulations). The data are subject of analysis and decision-making, in which statistical and probabilistic approaches or mathematical models are employed to develop the DTs. Throughout the entire process, various big-data processing tools that allow parallel processing on multiple nodes may be employed. The overall data flow for the development of a purpose-driven DTs begins with the creation of a data model. After the data model is developed and tested, data from both the physical and virtual twins are used to deploy specific services to achieve the given organizational objectives, such as design optimization, dynamic process planning, or content customization [25].

The presented architecture shows common ground with existing frameworks for specific domains; for example, the Open Systems Interconnection model consists of seven layers (physical, data link, network, transport, session, presentation, and application) and paves the way for layered network architectures with the use of abstraction layers. Therefore, further research on the reuse and unification of different reference architectures on DTs is required. Table 2 summarizes the layers, properties, and corresponding sources presented in this research.

Table 2. DTs reference architecture layers

| Layer | Properties | Sources, SLRs |
|--|--|--|
| 9. Business model, processes | Value propositions, revenue streams, cost structure, purpose | [6], [7], [21], [22], [23], [25], [27] |
| 8. System security, data privacy | Security, privacy, risks, abuse, encryption, authorization | [27] |
| 7. Micro-services, deployment | Micro-services, deployment, loose coupling, interfaces, applications | [7], [21], [23], [25] |
| 6. Data models, algorithms, virtual entity, virtual twin | Data models, algorithms, ML, decision rules | [7], [13], [19], [22], [24], [25], [27] |
| 5. Data preparation, data representation | Semantics, ontologies, data labels, fidelity, twinning accuracy | [5], [7], [13], [17], [19], [22], [25] |
| 4. Data storage, data integration | Databases, synchronization, big-data processing | [5], [7], [13], [17], [19], [22], [24], [25] |
| 3. Network, connectivity | Network architecture, data exchange protocols, middleware platforms, streaming | [5], [7], [13], [17], [19], [22], [24], [25] |
| 2. Data generation | Parameters, types of data, fidelity, accuracy | [5], [6], [7], [13], [17], [19], [22], [25] |
| 1. Physical entities, physical twin | Physical existence | [6], [7], [17], [19], [22], [25] |

3.4. Future research directions

Another objective of this meta-review is to summarize future research directions for DTs on the basis of the findings of the incorporated SLRs. The rapidly increasing DT popularity and scope, as well as the involvement of IoT, big data, and AI technologies, broadens the research challenges of digital twinning. We categorize these challenges in the following eight areas.

3.4.1. Concept development. Different application areas develop DT concepts from their specific domain.

Therefore, specified interactions and interfaces between various disciplines are required. This gives rise to the issue of a multidisciplinary development of the DT concept [5]. On the one hand, a general framework and reference architecture is required to develop DTs on a mutual background [13]. On the other hand, a broad array of domain-specific DTs will promote more detailed implementations in various application areas. With increasing research combining DTs with emerging technologies such as blockchain and virtual reality, applications in new fields such as infrastructure, education, and healthcare are imminent [7].

3.4.2. Business models. Research on the business pre-conditions and performance effects of DTs is relatively scarce. A wealth of research focuses on technical needs. In the future, other perspectives should also be considered, such as performance in terms of customer experience and business models. Some existing reviews and guidelines for requirements engineering could be a potential starting point for further research initiatives [5], [7].

3.4.3. Integration. With respect to the integration of DTs, technologies, devices, data, processes, and systems should be integrated together in a strategic and operating environment. The reviewed SLRs have addressed some integration issues, such as device integration, system integration, and data integration. However, full integration of DTs in organizational processes and enterprise systems has not yet been addressed [27]. Such systems can foster effective decisions, improve business processes, and make the enterprise more competitive. Therefore, the integration of DTs in enterprise systems is one of the main issues for successful implementation [5].

3.4.4. Data entry, data preparation, data augmentation. Further research needs to clarify several questions in the data-entry process, such as how the IoT facilitates data harvesting from a physical twin (using sensors), data integration, and data sharing with the corresponding virtual twins. This process can incur considerable costs. Sometimes, twinning can be more costly than the asset itself, in which case it makes little sense to create DTs. By contrast, the collected data are vast, heterogeneous in nature, unstructured, and noisy. Thus, further research on data pre-processing is required to ensure its effective use [21]. Specifically, applying data-cleaning techniques is necessary to organize and restructure data entry [5]. Furthermore, controlling the flow of such a large amount of data is a significant challenge. Finally, to improve the accuracy of the DT model, the underlying ML algorithms require a certain amount of data for training purposes [25]. The data acquisition problem is a significant challenge in the realization of DT models in small and medium-sized companies. Therefore, future research needs to consider

approaches for data incubators and data augmentation [21].

3.4.5. Big data. The explosive growth of social media and IoT technologies in the industrial sector has led to the generation of a large amount of data. To this end, big-data analytics requires advanced architectures, frameworks, technologies, tools, and algorithms to capture, store, share, process, and analyze the underlying data. There is also a potential for cloud- and edge-computing platforms to handle DT-related data. Cloud- and edge-computing integration allows DTs to process at a faster pace while processing vast amounts of heterogeneous and semantic data [7], [25]. Further research is required given the insufficient possibilities for synchronization between the physical and digital parts, the absence of high-fidelity models for simulation and virtual testing in different scales, the difficulties in predicting complex systems, and the challenges with gathering and processing large datasets [21].

3.4.6. Data analysis, ML, simulations. Algorithms for data analytics have played a major role in DTs for decision-making. However, the selection of a particular ML model with customized configuration is challenging. Every ML approach has diverse accuracy and efficiency levels with different applications and datasets. Therefore, depending on the motive and application of a DTs, the selection of the best ML algorithm and features is challenging [25]. Realizing modeling consistency and accuracy will improve the quality of DTs, enhancing the benefits of their applications [7]. Moreover, twinning processes might not only start from the physical entity but also be based on simulations at the level of the virtual entity. Therefore, different questions on data simulations are relevant for further research on DTs [24]. Another research topic pertains to the implementation of ML algorithms with respect to operations and continuous deployment. Implementation of an accurate multi-scale DT model of work-in-process scenarios is still challenging because the real-time changes during the twinning process are difficult to perceive and simulate [21].

3.4.7. Standardization. Although many DTs have been developed in various industries, the creation of a complex and reliable DT demands standardization. Currently, no single standard focuses solely on digital twinning. The ISO/DIS 23247-1 standard has limited information on digital twinning, and therefore DT deployment challenges are continuing to grow as a result of the lack of standardization [21]. Standardization efforts are underway by the joint advisory group (JAG) of ISO and IEC on emerging technologies [25]. Many specific architectures for DTs are proposed, but integrated design standards need to be investigated by designers who take all disciplines into

consideration simultaneously [5]. RAMI 4.0 provides such a holistic view of important aspects in smart manufacturing that different stakeholders need. It combines three core dimensions in a cuboidal space covering (1) the whole life cycle from development to disposal and resource recovery, (2) multi-layer integration from asset to business, and (3) the connection from products to the IoT and services.

3.4.8. Security and privacy. Some concepts such as human-, product-, energy-, city-, and defense-related DTs, are considered critical and may require stringent security and privacy regulations [21]. First, with the involvement of IoT devices in digital twinning, emphasis needs to be put on the security of the underlying communication protocols [27]. Second, the large collection of asset-related data needs to be stored securely, to prevent data breaches from insider and outsider threats [25].

4. Discussion

This paper deals with a meta-review of 14 SLRs on DTs published between 2018 and 2021. We analyze the results of the meta-review with respect to conceptualization, application areas, reference architecture, and future research directions for DTs.

An evaluation of results leads to several important implications for research and management practice on DTs. First, the conceptualization of DTs is coined by the area of its application. Early works on the subject focused solely on smart manufacturing. A general understanding of the concept defines a DT as a cyber-physical system with physical and digital parts. Data flows between a physical and digital object with full integration in both directions can be viewed as the central property of DTs [17]. Second, conceptualization of DTs needs to be extended by other properties such as data models, connectivity, accuracy, and synchronization [13], [24].

Application areas for DTs are expanding and cover a broad array of domains, from manufacturing and healthcare to smart cities, logistics, business, economics, and even psychology. However, a common reference architecture for DTs can define the relevant properties of DTs over several domains. Therefore, this paper presents a unified reference architecture with nine distinct layers as a blueprint for the configuration of DTs. Finally, the meta-review unpacks future research directions for DTs in eight different areas: concept development, business model, integration, data entry, big data, data analysis, standardization, and security and privacy. This opens pathways for future research and highlights the challenges for the further practical implementation of DTs.

5. References

- [1] 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.
- [2] F. Tao and M. Zhang, "Digital Twin Shop-Floor: A New Shop-Floor Paradigm Towards Smart Manufacturing," *IEEE Access*, vol. 5, pp. 20418–20427, 2017, doi: 10.1109/ACCESS.2017.2756069.
- [3] Q. Qi and F. Tao, "Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison," *IEEE Access*, vol. 6, pp. 3585–3593, 2018, doi: 10.1109/ACCESS.2018.2793265.
- [4] M. Grieves, "Digital twin: manufacturing excellence through virtual factory replication," *White Pap.*, vol. 1, pp. 1–7, 2014.
- [5] X. Wu, V. Goepp, and A. Siadat, "Concept and engineering development of cyber physical production systems: a systematic literature review," *Int. J. Adv. Manuf. Technol.*, vol. 111, no. 1–2, pp. 243–261, Nov. 2020, doi: 10.1007/s00170-020-06110-2.
- [6] H. Zhang, L. Ma, J. Sun, H. Lin, and M. Thürer, "Digital twin in services and industrial product service systems: Review and analysis," *Procedia CIRP*, vol. 83, pp. 57–60, 2019.
- [7] K. Y. H. Lim, P. Zheng, and C. H. Chen, "A state-of-the-art survey of Digital Twin: techniques, engineering product lifecycle management and business innovation perspectives," *Journal of Intelligent Manufacturing*, vol. 31, no. 6, Springer, pp. 1313–1337, 01-Aug-2020, doi: 10.1007/s10845-019-01512-w.
- [8] Y. Lu, C. Liu, K. I.-K. Wang, H. Huang, and X. Xu, "Digital Twin-driven smart manufacturing: Connotation, reference model, applications and research issues.," *Robot. Comput. Manuf.*, vol. 61, p. N.PAG-N.PAG, Feb. 2020.
- [9] N. Mohammadi and J. E. Taylor, "Smart city digital twins," in *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2017, pp. 1–5, doi: 10.1109/SSCI.2017.8285439.
- [10] Y. Liu *et al.*, "A Novel Cloud-Based Framework for the Elderly Healthcare Services Using Digital Twin," *IEEE Access*, vol. 7, pp. 49088–49101, 2019, doi: 10.1109/ACCESS.2019.2909828.
- [11] R. Parmar, A. Leiponen, and L. D. W. Thomas, "Building an organizational digital twin.," *Bus. Horiz.*, vol. 63, no. 6, pp. 725–736, Nov. 2020.
- [12] A. Gaggioli, "Digital Twins: An Emerging Paradigm in Cyberpsychology Research?," *CyberPsychology, Behav. Soc. Netw.*, vol. 21, no. 7, pp. 468–469, Jul. 2018.
- [13] H. van der Valk, H. Haße, F. Möller, M. Arbter, J.-L. Henning, and B. Otto, "A Taxonomy of Digital Twins," in *Proc. 26th Americas Conference on Information Systems*, 2020, pp. 1–10.
- [14] B. Kitchenham and S. Charters, "Guidelines for

- performing systematic literature reviews in software engineering,” *EBSE Tech. Rep. EBSE-2007-01*, 2007.
- [15] B. Kitchenham, O. P. Brereton, D. Budgen, M. Turner, J. Bailey, and S. Linkman, “Systematic literature reviews in software engineering—a systematic literature review,” *Inf. Softw. Technol.*, vol. 51, no. 1, pp. 7–15, 2009.
- [16] Z. M. Cinar, A. A. Nuhu, Q. Zeeshan, and O. Korhan, “Digital Twins for Industry 4.0: A Review,” in *Global Joint Conference on Industrial Engineering and Its Application Areas*, 2019, pp. 193–203.
- [17] W. Kritzing, M. Karner, G. Traar, J. Henjes, and W. Sihm, “Digital Twin in manufacturing: A categorical literature review and classification,” *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1016–1022, 2018.
- [18] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, and F. Sui, “Digital twin-driven product design, manufacturing and service with big data,” *Int. J. Adv. Manuf. Technol.*, vol. 94, no. 9–12, pp. 3563–3576, 2018, doi: 10.1007/s00170-017-0233-1.
- [19] T. Catarci, D. Firmani, F. Leotta, F. Mandreoli, M. Mecella, and F. Sapio, “A Conceptual Architecture and Model for Smart Manufacturing Relying on Service-Based Digital Twins,” in *2019 IEEE International Conference on Web Services (ICWS)*, 2019, pp. 229–236, doi: 10.1109/ICWS.2019.00047.
- [20] K. Y. H. Lim, P. Zheng, and C.-H. Chen, “A state-of-the-art survey of Digital Twin: techniques, engineering product lifecycle management and business innovation perspectives.,” *J. Intell. Manuf.*, vol. 31, no. 6, pp. 1313–1337, Aug. 2020.
- [21] T. Y. Melesse, V. Di Pasquale, and S. Riemma, “Digital twin models in industrial operations: A systematic literature review,” *Procedia Manuf.*, vol. 42, pp. 267–272, 2020.
- [22] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, “Characterising the Digital Twin: A systematic literature review,” *CIRP J. Manuf. Sci. Technol.*, vol. 29, pp. 36–52, 2020.
- [23] M. R. Enders and N. Hoßbach, “Dimensions of digital twin applications—a literature review,” 2019.
- [24] C. H. dos Santos, J. A. B. Montevechi, J. A. de Queiroz, R. de Carvalho Miranda, and F. Leal, “Decision support in productive processes through DES and ABS in the Digital Twin era: a systematic literature review,” *International Journal of Production Research*. Taylor and Francis Ltd., 2021, doi: 10.1080/00207543.2021.1898691.
- [25] M. M. Rathore, S. A. Shah, D. Shukla, E. Bentafat, and S. Bakiras, “The Role of AI, Machine Learning, and Big Data in Digital Twinning: A Systematic Literature Review, Challenges, and Opportunities,” *IEEE Access*, vol. 9, pp. 32030–32052, 2021.
- [26] B. Ketzler, V. Naserentin, F. Latino, C. Zangelidis, L. Thuvander, and A. Logg, “Digital Twins for Cities: A State of the Art Review,” *Built Environ.*, vol. 46, no. 4, pp. 547–573, 2020.
- [27] M. Perno, L. Hvam, and A. Haug, “Enablers and Barriers to the Implementation of Digital Twins in the Process Industry: A Systematic Literature Review,” in *2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2020, pp. 959–964, doi: 10.1109/IEEM45057.2020.9309745.
- [28] K. Y. H. Lim, P. Zheng, and C.-H. Chen, “A state-of-the-art survey of Digital Twin: techniques, engineering product lifecycle management and business innovation perspectives,” *J. Intell. Manuf.*, pp. 1–25, 2019.
- [29] R. Soderberg, K. Warmefjord, J. S. Carlson, and L. Lindkvist, “Toward a Digital Twin for real-time geometry assurance in individualized production,” *Cirp Ann. Technol.*, vol. 66, no. 1, pp. 137–140, 2017, doi: 10.1016/j.cirp.2017.04.038.
- [30] B. A. Talkhestani, N. Jazdi, W. Schloegl, and M. Weyrich, “Consistency check to synchronize the Digital Twin of manufacturing automation based on anchor points,” in *51st Cirp Conference on Manufacturing Systems*, vol. 72, L. Wang, Ed. 2018, pp. 159–164.
- [31] J. Thönes, “Microservices,” *IEEE Softw.*, vol. 32, no. 1, p. 116, 2015.
- [32] A. Rojko, “Industry 4.0 concept: Background and overview.,” *Int. J. Interact. Mob. Technol.*, vol. 11, no. 5, 2017.
- [33] F. Adrodegari, N. Saccani, C. Kowalkowski, and J. Vilo, “PSS business model conceptualization and application,” *Prod. Plan. Control*, vol. 28, no. 15, pp. 1251–1263, 2017.
- [34] M. Ghobakhloo, “The future of manufacturing industry: a strategic roadmap toward Industry 4.0,” *J. Manuf. Technol. Manag.*, 2018.