

Digital Twins as Software and Service Development Ecosystems in Industry 4.0: Towards A Research Agenda

Yueqiang Xu ¹[0000-0002-3528-423X], Tero Päivärinta ²[0000-0002-7477-0783] and Pasi Kuvaja ²[0000-0002-1488-6928]

¹ Martti Ahtisaari Institute

² M3S, Empirical Software Engineering on Software, Systems, and Services
Faculty of Information Technology and Electrical Engineering

^{1,2} Pentti Kaiteran katu 1, FI-90014 University of Oulu, Finland
{yueqiang.xu, tero.paivarinta, pasi.kuvaja}@oulu.fi

Abstract. While research on digital twins of cyber-physical systems within industry 4.0 is emerging, the software development perspective on digital twins remains under-explored. Contemporary definitions and examples of digital twins have covered company- or product-specific solutions or discussed the use of digital twins in rather proprietary value chains. This paper addresses the importance of taking an ecosystem view on software development on digital twins for industry 4.0 and outlines a framework for building a research agenda for such ecosystems. The framework includes three dimensions: scope of the digital twin software platform (internal, value chain, ecosystem), life-cycle phases of the industry 4.0 system related with the digital twin (creation, production, operation & maintenance, disposal), and level of integration between the twin and the physical system (model, shadow, twin). As this research-in-progress addresses examples of research questions in light of the framework, further research to build a full-scale research agenda based on a systematic literature review is suggested.

Keywords: digital twin, software platform, service ecosystem, research agenda, industry 4.0

1 Introduction

Industry 4.0 has attracted growing scholarly attention in recent years [1]. Progress in information technologies (IT), big data, and intelligent manufacturing have stimulated and enabled the global transformation in the manufacturing industry [2]. With the recent advances in the Internet of Things (IoT) in the modern industry, the conventional industrial production planning and management paradigm demands a shift or upgrade from purely providing surveillance-centric functions to building a comprehensive information framework of the industrial processes [3].

In this vein, an emerging concept of *digital twin* has gained momentum and much attention [4]. The term was coined by Michael Grieves in 2002 as a virtual, digital equivalent of a physical product [5] in context of an executive course on Product

Lifecycle Management (PLM) [6]. A digital twin is “a set of virtual information constructs that fully describes a potential or actual physical manufactured product” [6, p. 94]. Moreover, a digital twin can be defined at two levels: *digital twin prototype (DTP)*, which contains digital information “necessary to describe and produce a physical version that duplicates or twins the virtual version”, and *digital twin instance (DTI)*, which “describes a specific corresponding physical product that an individual Digital Twin remains linked to throughout the life of that physical product” [6, p. 94].

From the viewpoint of cyber-physical systems (CPSs), a digital twin is defined as a specific applied technical framework as a realization of a CPS [7], while from the product life cycle perspective, a digital twin is a dynamic digital replica of physical assets, processes, and systems that comprehensively monitors the entire product lifecycle [3]. Grieves & Vickers propose also the concept of *Digital Twin Environment (DTE)* to cover an integrated multi-domain “application space for operating on Digital Twins for a variety of purposes” [6, p. 94]. Digital twin's backbone involves IoT technologies for real-time and multisource data acquisition. Additionally, artificial intelligence (AI) and software analysis are incorporated as part of the system to create digital simulation models that are dynamically updated and modified with their physical counterparts. The digital twin also advances the data visualization schemes, such as Virtual Reality (VR) and Augmented Reality (AR), so as to create more vivid, seamless, and user-friendly interfaces and user experience, for instance, for multimodal interactions with the physical system in question [8].

Overall, research (by Autumn 2019) on digital twins is emerging, diverse, scattered and weakly-linked. For example, there is a need for research in relation to the challenges of defining the unified framework for building the digital twin at the systemic level [9] while others call for the modular approach and continue to work on optimizing the synchronization framework between the cyber and physical worlds [10]. A majority of the literature has been primarily focused on the design schemes of the digital twin in narrow technical applications (e.g. surveillance signal processing [3], multimodal user interface adaptation [11]), while [12] suggests that a (technological) system needs a coherent view at different levels or aspects such as core technology, platform, and application levels. Last but not least, the software development and implementation perspective on digital twin remains an under-explored territory in the existing studies. We argue that the development perspective on digital twins as software platforms within adjacent service development ecosystems is a crucial centerpiece to facilitate and enable the proliferation of digital twin applications in industry 4.0.

In the domain of Industry 4.0, asset administration related to an industrial plant and related CPSs requires multiple stakeholders throughout the plant life cycle, such as device manufacturers, system integrators, plant owners and Industry 4.0 architects (e.g., [13]). While architectural standards are still emerging, digital twins are expected to change current business models, development approaches, and technologies, while posing significant challenges to the architects and stakeholders in need to co-operate in this domain (e.g., [13], [14]). This trend highlights that development of innovative services and software applications on digital twins (such as those targeted to diagnostics, predictive maintenance, training, or virtual support for fieldwork on industrial artifacts) is

increasingly taking place in platforms of actors in a value chain or even ecosystems of actors throughout the CPS life-cycle.

In an effort to understand the development of digital twins in industry 4.0, this paper outlines the challenges and proposes a framework for building a research agenda for the development of digital twins with related software and services at the platform and ecosystem levels. This paper sheds light on positioning digital twin research challenges related to the ecosystem and platform views on digital twins in the manufacturing industry.

The rest of this paper is organized as follows. Section 2 provides the review and discussion on the definition of the digital twin in the literature. Section 3 introduces the concepts of platform and ecosystem as an overarching framework to systematically identify the challenges and research agenda concerning the software development of a digital twin, especially in the ecosystem setting. Section 4 focuses on the presentation and proposition of the key research agenda for digital twins. Section 5 summarizes the contributions of our work and concludes the paper.

2 Analytical Dimensions of Digital Twins

2.1 High Level Structure of Digital Twins

The fundamental idea of the digital twin includes the elements of the continuous real-time data exchange between the physical space, e.g. smart factory and the digital space, the data-driven training loop for continuous optimization and improvement of production as well as the integration of production and information elements [7]. A high-level illustration of digital twin is presented in Fig. 1.

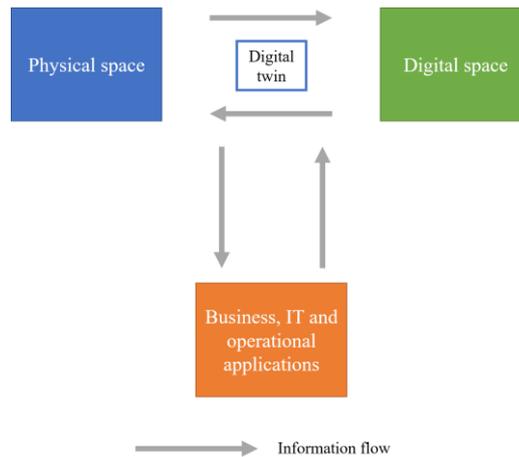


Fig. 1. High-level illustration of digital twin (adapted from [15])

Two key aspects of the digital twin system architecture are: 1) a tight logical connection between the digital twins, the physical world counterparts and between different digital twins that are needed to realize a dedicated system functionality; 2) a changeable logical system architecture. Digital twin components may change properties like interfaces, connections to other components, behavior and simulation models, as well as communication protocols. Therefore, the initial logical architecture (of the DTP) may differ from the actual logical architecture (of a particular DTI) during runtime [14].

2.2 PLM Stages and Digital Twins

Based on the existing literature discussing digital twin in PLM [16], digital twin's integration into the PLM can include the following tasks and processes (at a coarse level): Creation, Production, Operations (and Maintenance), and Disposal [6].

In the Creation phase, a digital twin can be connected to design drawings, behavior prediction, design simulation, engineering data, manufacturing operations management key performance indicators (KPIs) and design tools. The digital twin can include order planning based on statistical assumptions, improved decision support by means of detailed diagnosis as well as automatic planning and execution from orders by the production units [17].

In the Production phase, the digital twin can focus on such issues as manufacturing instructions, virtual commissioning, and Hardware (HW)-in-the-loop tests [18].

In the Operations and Maintenance phase, digital twins can provide operation instructions, plant visualization, control performance, what-if studies, operational data, real-time movement, operational KPIs, and operational state display. The digital twin can be used to identify the impact of state changes on upstream and downstream processes of a production system [19, 20].

In maintenance actions, involved in the operations phase, digital twins are used to diagnose the system run-time, identification, and evaluation of anticipatory maintenance measures, the evaluation of machine conditions based on descriptive methods and machine learning algorithms [21]. As well, a digital twin can facilitate the integration, management, and analysis of machinery or process data during different stages of the machine life cycle to handle data/information more efficiently and further achieve better transparency of a machine's health condition [22].

The disposal phase has been so far less prominent in the mainstream discussion on digital twins. However, digital twin solutions within and after the disposal phase of the physical system life-cycle may appear important as well – for example, in the field of nuclear plants and spent nuclear fuel [6].

2.3 Three Levels of Digital Twin Integration

Digital twins were initially considered to be high fidelity mathematical models to reflect the behavior of the actual system as close as possible [18, 23]; this perception has evolved to include simulated and visible dynamic 3D models of real-world assets. The digital twin definition has been enriched over the time to be an evolving digital profile

of the historical and current behavior and all properties of an asset, where an asset can be anything of value for an organization such as a physical device, a subsystem, a plant, or a software entity.

Based on the given definitions of a digital twin in any context, digital twins represent digital counterparts of physical objects. It is argued by [21] that three levels of how such counterparts interact are interchangeably used in the literature, namely, digital model, digital shadow and digital twin. Evidently, some digital representations are modeled manually and are not connected with any physical object in existence. At the same time, other digital twins are fully integrated with real-time data exchange. Thus, the definition of digital twins needs to be classified at different levels of how the digital counterparts interact with the physical objects to endorse a clear understanding of the existing and future research.

Based on the categorical classification of the digital twin [21], a digital model is a digital representation of an existing or planned physical object that involves no form of automated data exchange between the physical object and the digital object. Digital data of existing physical systems might still be in use for the development of such models, but all data exchange is done manually. A change in the state of the physical object has no direct effect on the digital object and vice versa as illustrated in Fig. 2.

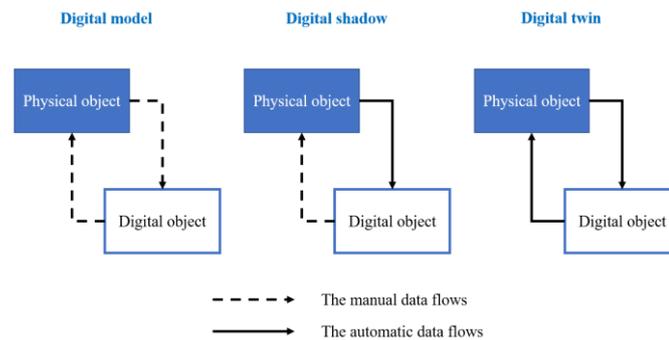


Fig. 2. Digital model, digital shadow, and digital twin [21]

The digital shadow utilizes automated one-way data flow between the state of an existing physical object and a digital object, becoming a virtual “shadow” of the physical object or system. A change in the state of the physical object or system changes the modelled state in the digital object [21]. In a full digital twin, the data flow between an existing physical object and a digital object is fully integrated to both directions. The digital object or system controls or affects a physical object. There might also be other physical or digital objects, which can make changes of state in the digital object. A change in the state of the physical object directly leads to a change in the state of the digital object, and this process goes in both directions as can be seen in Fig. 2.

3 Digital Twin Platforms and Ecosystems

3.1 Digital Platforms

When developing a variety of products that can meet the needs of a large number of customers, the manufacturers strive for commonalities and distinguishing features of the developed product variants. A common approach for this is the adoption of a platform approach [24], which is typically achieved by modularizing the product's architecture [25] or the production system [26]. The key to a platform approach is to achieve economies of scale in production [27] by using common components for a variety of product configurations [28].

In the literature, there are numerous platform types with different definitions and properties, e.g. the two-sided platform of [29]. To facilitate and navigate the literature of platform, the study of [30] divides the discussion of the digital platform into two perspectives: economics and engineering design. From an economic point of view, digital platforms are seen as markets where the platform facilitates exchanges between actors who otherwise would not be able to do business with each other. An example of this type of platform can be Airbnb, as the platform connects the business and leisure travelers with homeowners and enables the transaction between the buyers and sellers from different sides of the platform.

3.2 Digital Twin as Software Platform and Ecosystem

The technological perspective of the platform considers that a platform is a technological architecture that enables innovation [30, 31]. One of the basic ideas is that the platform inherently has the modularized features, functions, and components that can be used to decompose complex systems into manageable components connected by interfaces. Based on the level of authorization, the different modules are accessed and share data and information on the platform [28]. Furthermore, it is categorized by [30] the platforms into three organizational categories: the internal platform, the supply chain platform and the industry ecosystem platform (Table 1).

Table 1. Integrative classification of technological platforms (adapted from [30]).

	Internal platform	Supply-chain platform	Ecosystem platform
Unit of analysis	Company	Supply chain	Industry ecosystem
Degree of openness	Closed interfaces	Selectively open and closed interfaces	Open interfaces
The level of access	Company capabilities	Supply-chain capabilities	Unlimited number of external sources
Control/governance mechanisms	Managerial hierarchy	Contracts between supply chain companies	Ecosystem governance
Literature	[30, 32]	[30, 33, 34]	[30, 35, 36]

Digital Twin Platforms, Units of Analysis. Gawer [30] suggests the following units of analyses while looking at software and service platforms. An organization's *internal* platform often stands for the enterprise information system are sole available and accessible within a company or organization (e.g. an internal Enterprise resource planning (ERP) system of a factory), while the *supply chain* platform and *ecosystem* platform are outside a focal organization and exist in a broader network or ecosystem. As such, each platform type can be considered as a different unit of analysis. The internal platform only defines a single company/organization as the unit of analysis, while the boundary of an ecosystem platform covers multiple interconnected and interdependent actors with an ecosystem. [30].

Digital Twin Platforms, Degree of Openness. The degree of openness is a defining characteristic of a platform. The traditional top-down enterprise management systems operate in a closed and hierarchical manner with higher-level controllers to plan, coordinate, and optimize production processes start to shift to changes that enable greater flexibility and adaptability to internal and external changes. In the world of networked production with multiple actors in complex manufacturing and production processes and ecosystems, the control process will inherently be a hybrid combination of hierarchical and heterarchical collaboration processes [15].

The equipment and technology vendors and software and system developers are all needed to become part of the ever-increasing complexity of larger and larger automated production networks [37], such as a digital twin network or ecosystem to support the production systems in industry 4.0 [15]. In this light, interconnectivity plays a crucial role not only within the boundaries of the manufacturing company in order to optimize production processes, but also beyond these limits. As a result, manufacturing is inherently a distributed control process [15].

A digital twin can improve the transparency of the production process beyond the organizational boundaries of the manufacturing company and data supports business applications on any device, anytime, anywhere [15]. The data-intensive nature of smart factory systems allows for timely, accurate, and detailed logging traces that enable real-time viewing of many systems and activities that were previously unavailable [15]. Evidently, today's production systems are no longer fully closed as [30]'s internal platforms.

Access to Capabilities and Resources. The platform's openness does not only affect actors' access to information, naturally, it also influences the platform's access to capabilities and resources as an increasing number of participators means a larger number of source pool and stronger capabilities, such as the open innovation capability in the literature [38, 39].

By incorporating the open innovation and ecosystem literature, we elaborate on [30]'s original proposition that focuses solely on the platform's innovation capabilities. We enrich this conceptualization and propose that platforms (especially in the ecosystem setting) can offer varying levels of access to resources and capabilities. Therefore, in contrast to a closed, internal platform focusing on internal development, a more open

ecosystem platform supports also the external actors to access, exchange and share resources (e.g., software, information and data) and capabilities (e.g., domain-based technical expertise) through the platform [30].

Platform Control and Governance. Orchestration and choreography [40] are two well-known composition strategies in the area of service-oriented architectures to create business processes from individual web services. According to [15], these concepts are partly related to the hierarchical and heterarchical functioning of manufacturing control systems. First, orchestration interacts with internal and external web services to take control from a stakeholder perspective. On the other hand, the choreography reflects a more collaborative interaction in which each involved stakeholder can describe his share of the interaction [15].

It is suggested by [41] that virtual factories that combine service-oriented computing and service-oriented workflows with the IoT. Similar to the typical roles of a service requester, a service provider, and a service register in a service-oriented architecture, they propose service providers, service consumers, and virtual factory brokers as essential roles of the virtual factory. The latter role is responsible for managing and controlling the virtual factory and uses services to model manufacturing processes and assemble products based on the results of factories of various business partners. The plug-and-play character of digital twin and virtual factories helps companies execute cross-organizational manufacturing processes as if they were running in a single company. Overall, as an emerging concept, the digital twin's technological development is at its nascent stage. Multiple aspects of the digital twin platform still require further development. This paper will focus more on software development with the platform and ecosystem perspective in the next section.

4 Framework for Building a Research Agenda

To outline existing (recognized) research issues and gaps, we constructed a three-dimensional framework based on the above-mentioned, three streams of literature related to the scope of the industrial organization and structure [30], the life-cycle perspective of the digital twin in industry 4.0 [6], and the levels of integration of digital twin [21].

First, **the scope dimension** consists of three organizational scopes: internal operations of a particular organization, the value chain of a group of suppliers, producers and distributors, and the ecosystem that includes organizations beyond a particular value chain (see Fig. 3).

Second, we utilize the product **life-cycle perspective** on digital twins [6] as the second dimension. Essentially, the life-cycle dimension defines the four stages of the industrial 4.0 processes: creation and design of the products, the production process, operation and maintenance process, and the disposal of the wastes, defects, and decommissioned products (Fig. 3).

Third, we employ the categorization on the **levels of integration** between digital twins and the represented physical objects [21] as the third dimension of the framework

(Fig. 3). The three levels of integration include the aforementioned concepts of the digital model, the digital shadow, and the digital twin. We use acronyms of DM (digital model), DS (digital shadow), and DT (digital twin), respectively, in Fig. 4. The level of integration dimension can be observed in each cell of the framework to assist the classification of the research issues and gaps concerning the digital twin studies.

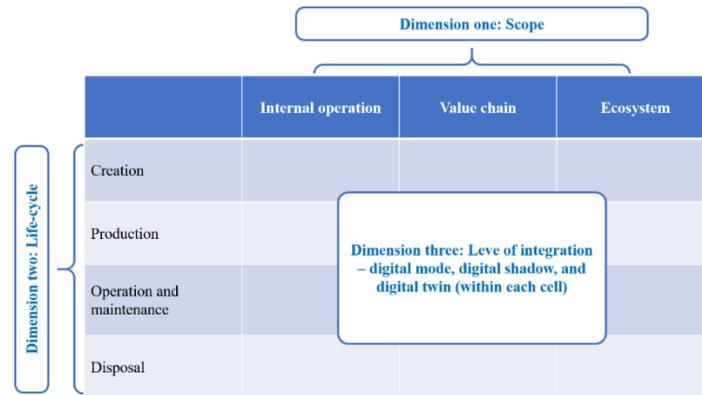


Fig. 3. Research agenda framework

By utilizing the three-dimensional framework as the conceptual lens, we mapped a number of research gaps and challenges that dwell in the digital twin research domain. The overall results are presented in Fig. 4.

Internal Digital Twin. Digital twin research focusing on internal operations of an organization has specific technical issues to address in different stages of the production life-cycle. Data management and integration challenges from the viewpoint of the whole CPS to be twinned within one host organization is an important issue to be addressed across the internal production life-cycle, especially in creation, production, and operation and maintenance stages.

Data capture from IoT represents challenges in such stages as production and operation and maintenance. For example, in certain industrial contexts, digital twin utilizes high-density intensive network cameras to ensure seamless monitoring. On the one hand, when processing intensively networked video, the computer architecture must be updated to meet the specific use case requirements. On the other hand, enabling a resource-constrained IoT device with modern analysis techniques, e.g., deep learning, can help relieve the pressure of cloud infrastructure and save the network bandwidth [3]. Another challenge lies within the internal organization is in regard to understanding needs and opportunities for data utilization in the later phases of the CPS life-cycle of a complex CPS, for instance, how predictive diagnosis can be utilized in operation and maintenance stage and even disposal stage. In general, the extant research in digital twin has made the process from the conventional digital model to the digital shadow and digital twin. However, the study identifies the fully digital twin development remains a gap to be addressed by the research community.

Dimension three: Level of integration (DM, DS, DT)		Dimension one: Scope		
		Internal operation	Value chain	Ecosystem
Dimension two: Life-cycle	Creation	<ul style="list-style-type: none"> Data management and integration (DS, DT) 	<ul style="list-style-type: none"> Data governance (DM, DS, DT) 	<ul style="list-style-type: none"> Ecosystem orchestration and governance (DM, DS, DT) The degree of openness (DM, DS, DT) Data provisioning and access (DM, DS, DT)
	Production	<ul style="list-style-type: none"> Data capture from IoT (DS, DT) Data management and integration (DS, DT) 	<ul style="list-style-type: none"> Data interoperability (DM, DS, DT) Data governance (DM, DS, DT) Integration of intelligence (DS, DT) 	<ul style="list-style-type: none"> Ecosystem orchestration and governance (DM, DS, DT) The degree of openness (DM, DS, DT) Data provisioning and access (DM, DS, DT)
	Operation and maintenance	<ul style="list-style-type: none"> Understanding needs and opportunities for data utilization (DS, DT) Data management and integration (DS, DT) Understanding needs and opportunities for data utilization (DM, DS, DT) 	<ul style="list-style-type: none"> Data interoperability (DM, DS, DT) Data governance (DM, DS, DT) Integration of intelligence (DS, DT) 	<ul style="list-style-type: none"> Ecosystem orchestration and governance (DM, DS, DT) The degree of openness (DM, DS, DT) Data provisioning and access (DM, DS, DT)
	Disposal	<ul style="list-style-type: none"> Understanding needs and opportunities for data utilization (DM, DS, DT) 	<ul style="list-style-type: none"> Data governance (DM, DS, DT) 	<ul style="list-style-type: none"> Ecosystem orchestration and governance (DM, DS, DT)

Fig. 4. Examples of identified research issues in light of the framework

Value Chain -Focused Digital Twin. When we expand the digital twin scope from internal organization to the level of a value chain, new challenges regarding data interoperability and data governance emerge.

First, data interoperability challenges are related to enhancing data capture by component providers, for instance, how to standardize data acquisition and exchange within a value chain. Second, data governance challenges arise as the organizational boundary of digital twin shifts from internal operations to a value chain. As an example, the intellectual property rights (IPRs) within the value chain with regard to data analytics and processing is an important issue. Third, the integration of intelligence (e.g. AI and machine learning) in the value chain is a potential area for software or web applications to explore further. Supporting the integration of the AI and machine learning capabilities can resolve the resource constraints that exist in contemporary digital twin systems. Complexity increases when digital twins exit in the much wider application scope. Large technology companies like Google have been focusing on tackling the challenges and issues of data analytics integration. For example, Google has enabled TensorFlow (Google's machine learning and deep learning software library) to be able to be implemented in the mobile device through JavaScript, which solves the resource

constrains for AI and machine learning integration in different systems and devices. In the context of the value chain, all three levels of integration (digital model, digital shadow, and digital twin) demand further research efforts.

Ecosystem-Focused Digital Twin. Currently, the digital twin ecosystem is primarily defined as a technical architecture that integrates the modeling and simulation while taking advantage of the industrial Internet of Things for data acquisition and information processing with cloud computing [42]. The technical ecosystem-focused digital twin system is visible and expected to be integrated into all phases of the PLM. The key challenges can be found in the areas of ecosystem governance mechanisms, degree of openness, and data sharing and access. The overall concept of suggesting digital twins as a basis for software and service development ecosystems represents a new contextual scope and view on related research. Such a view contributes to the existing digital twin research that mainly targets internal operations or value chain. Such a view will need to cover all the four stages of the CPS / product life-cycles.

First, the governance mechanisms concerning software ecosystem orchestration for digital twins requires further investigation. Our research has so far identified little or no discussion of digital twins related to platform thinking in the context of industry 4.0 beyond such platform concepts as digital twin clouds [43]. Challenges arise when the traditional industry moves from the closed internal platforms to the more open ecosystems.

As suggested by [44], digital twin components must interact and work together to solve complex problems. Digital twins encompass three types of interaction and collaboration: physical-to-physical, virtual-to-virtual, and virtual-to-physical. Through physical-to-physical interaction, multiple physical entities can communicate, coordinate, and collaborate to perform a complex task that cannot be performed by any single entity. Through virtual-to-virtual interaction, multiple virtual models can be connected to a network for information sharing. Through virtual-to-physical interaction and collaboration, the virtual model can be tuned in synchronization with the physical object, while the physical object can be dynamically adjusted based on direct commands from the virtual model.

Second, the degree of openness regarding the shift from proprietary software to open-source software represents another research gap in relation to digital twin ecosystems for Industry 4.0. In relation to digital twin platforms, the majority of previous literature deals exclusively with challenges related to technical development. For example, it is argued by [45] that integrating and interoperating systems is a major difficulty. These technologies are expensive and complex, making the use of digital twins more difficult [45]. However, numerous existing industrial software applications are proprietary, which creates barriers for industry players to embrace open-source software. There is a dilemma faced by the platform owner (particularly the industry incumbents) to protect proprietary intellectual properties while harnessing the scope and scale advantages of integrated, open-source software development environments.

Third, the data access related to data provisioning, utilization, and sharing is another under-explored area. As discussed in the previous section, the data fusion of digital twins involves three processes: data preprocessing, data mining and data optimization.

First, digital twins have to process a huge amount of data, including physical data, virtual data, and fusion data between them. Therefore, a data preprocessing that includes data cleansing, data conversion, and data filtering must be performed. Second, the preprocessed data is determined by fuzzy sets, rule-based closure, intelligent algorithms, and other advanced data analysis methods. Third, theories of data optimization are useful for dealing with the iterations of physical data, virtual data, connection data, service data, and data fusion, to discover the data evolution laws [44].

In the course of digitization, data on both products and processes are increasingly being created in the process industry. The data is needed to create digital twins, and more data is generated when digital twins are applied while code is being developed and analyzed. However, the study of [28] clearly shows that the digital twin ecosystem actors are unwilling to share both product and process-specific data, as well as code-specific data. In terms of product and process-specific data, end-users and machine builders are concerned that their data will be passed on to competitors, thereby revealing their company-specific secrets. Digital twins do not minimize these concerns because they need detailed information about products and processes in order to be fully exploited.

5 Conclusion and Future Work

Digital platforms are transforming nearly every industry today [36] revolutionizing future value creation in the manufacturing industry as they enable companies to become more digitalized [46]. The advent of industry 4.0 fosters digital platforms and the concept of a digital twin to attract scholarly attention and interest as the industries start to realize and understand related technological and business opportunities [47].

Digital twin as an emerging concept represents new research issues crossing the disciplines of industrial engineering, software engineering, cyber-physical systems, AI and machine learning, and ecosystem and platforms. Digital twins need to be investigated holistically throughout three organizational scopes (internal operations, value chain, and ecosystem) while the challenges of interoperability through the combination of completely different models, systems and tools for digital twin become paramount. From the product and CPS life-cycle perspective [6], the digital twin must be seamlessly integrated with the existing systems, processes, and models [9] in different life-cycle stages to realize its true potential in industry 4.0. Additionally, this study incorporates the levels of integration for digital twin [21] to decrease the conceptual ambiguity in the existing digital twin research. New challenges and research issues are expected to emerge as novel applications on digital twins are required for system life-cycle phases of operations, maintenance, and disposal; built upon expanding software platforms opened for expanding digital service ecosystems, and involving increasingly integrated full-fledged twins instead of mere models or shadows of the physical systems. The key contributions of this study are constructing a three-dimensional framework to explore and analyze the research challenges and gaps in a structured manner and investigating the ecosystem context for digital twin that is rarely addressed in the digital twin studies. Further studies can utilize the framework to systematically study

the digital twin research in both theoretical and empirical contexts, in order to spot knowledge gaps and new research issues for the future research agenda.

References

1. Uhlemann, T.H.J., Lehmann, C., Steinhilper, R.: The Digital Twin: Realizing the Cyber-Physical Production System for Industry 4.0. *Procedia CIRP*. 61, 335–340 (2017).
2. Qi, Q., Tao, F.: Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison. *IEEE Access*. 6, 3585–3593 (2018).
3. He, Y., Guo, J., Zheng, X.: From Surveillance to Digital Twin: Challenges and Recent Advances of Signal Processing for Industrial Internet of Things. *IEEE Signal Process. Mag.* 35, 120–129 (2018).
4. Enders, M.R., Hoßbach, N.: Dimensions of Digital Twin Applications - A Literature Review. *Twenty-fifth Am. Conf. Inf. Syst. Cancun, 2019*. 1–10 (2019).
5. Grieves, M.W.: Virtually Indistinguishable. In: *IFIP International Conference on Product Lifecycle Management*. pp. 226–242 (2012).
6. Grieves, M., Vickers, J.: Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In: *Transdisciplinary perspectives on complex systems*. pp. 85–113. Springer (2017).
7. Min, Q., Lu, Y., Liu, Z., Su, C., Wang, B.: Machine Learning based Digital Twin Framework for Production Optimization in Petrochemical Industry. *Int. J. Inf. Manage.* 1–18 (2019).
8. Saddik, A. El, Badawi, H., Alejandro, R., Velazquez, M., Laamarti, F.: IEEE COMSOC MMTC Communications -Frontiers Dtwins : A Digital Twins Ecosystem For Health And Well-Being IEEE COMSOC MMTC Communications - Frontiers Dtwins : A Digital Twins Ecosystem For Health And Well-Being. (2019).
9. Koulamas, C., Kalogeras, A.: Systems and Digital Twins in the Industrial Internet of Things. 95–98 (2018).
10. Kuts, V., Otto, T., Tähemaa, T., Bondarenko, Y.: digital twin based synchronised control and simulation of the industrial robotic cell using virtual reality. *Stomatopharyngology*. 26, 279–281 (2013).
11. B, K.J., Yigitbas, E., Engels, G.: A Digital Twin-Based Multi-modal UI Adaptation Framework for Assistance Systems in Industry 4 . 0. Springer International Publishing (2019).
12. Quan, X.I., Sanderson, J.: Understanding the artificial intelligence business ecosystem. *IEEE Eng. Manag. Rev.* 46, 22–25 (2018).
13. Wagner, C., Grothoff, J., Epple, U., Drath, R., Malakuti, S., Grüner, S., Hoffmeister, M., Zimmermann, P.: The role of the Industry 4.0 asset administration shell and the digital twin during the life cycle of a plant. In: *2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*. pp. 1–8 (2017).
14. Bauer, T., Antonino, P.O., Kuhn, T.: Towards Architecting Digital Twin-pervaded Systems. In: *Proceedings of the 7th International Workshop on Software Engineering for Systems-of-Systems and 13th Workshop on Distributed Software Development, Software Ecosystems and Systems-of-Systems*. pp. 66–69. IEEE Press, Piscataway, NJ,

- USA (2019).
15. Preuveneers, D., Ilie-zudor, E.: Robust digital twin compositions for Industry 4.0 smart manufacturing systems. 2018 IEEE 22nd Int. Enterp. Distrib. Object Comput. Work. 69–78 (2018).
 16. Arianfar, S., Kallenbach, J., Mitts, H., Mäkinen, O.: Back to the Future – Prediction of Incremental and Disruptive Innovations. 33–63.
 17. Boschert, S., Rosen, R.: Digital Twin---The Simulation Aspect. In: Hehenberger, P. and Bradley, D. (eds.) *Mechatronic Futures: Challenges and Solutions for Mechatronic Systems and their Designers*. pp. 59–74. Springer International Publishing, Cham (2016).
 18. Malakuti, S., Schlake, J., Ganz, C., Harper, K.E., Petersen, H.: Digital Twin: An Enabler for New Business Models.
 19. D’Addona, D.M., Ullah, A.M.M.S., Matarazzo, D.: Tool-wear prediction and pattern-recognition using artificial neural network and DNA-based computing. *J. Intell. Manuf.* 28, 1285–1301 (2017).
 20. Susto, G.A., Schirru, A., Pampuri, S., McLoone, S., Beghi, A.: Machine learning for predictive maintenance: A multiple classifier approach. *IEEE Trans. Ind. Informatics*. 11, 812–820 (2014).
 21. Kritzinger, W., Karner, M., Traar, G., Henjes, J., Sihn, W.: Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*. 51, 1016–1022 (2018).
 22. Lee, J., Lapira, E., Bagheri, B., Kao, H.: Recent advances and trends in predictive manufacturing systems in big data environment. *Manuf. Lett.* 1, 38–41 (2013).
 23. Brathall, L.G., Van Der Geest, R., Hofmann, H., Jellum, E., Korendo, Z., Martinez, R., Orkisz, M., Zeidler, C., Andersson, J.S.: Integrating hundred’s of products through one architecture-The industrial IT architecture. *Proceedings-International Conf. Softw. Eng.* 604–614 (2002).
 24. Meyer, M.H., Lehnerd, A.P.: *The Power of Product Platforms*. Free Press, New York (1997).
 25. Erixona, G., Anders, Y., Arnström, A.: Modularity – the Basis for Product and Factory Reengineering. *CIRP Ann. - Manuf. Technol.* 45, 1–6 (1996).
 26. Rogers, G.G., Bottaci, L.: Modular production systems: a new manufacturing paradigm. *J. Intell. Manuf.* 8, 147–156 (1997).
 27. Meyer, M.H., Osiyevskyy, O., Libaers, D., van Hugten, M.: Does product platforming pay off? *J. Prod. Innov. Manag.* 35, 66–87 (2018).
 28. Andersson, E.: *Enabling Successful Collaboration on Digital Platforms in the Manufacturing Industry A Study of Digital Twins*. (2019).
 29. Rochet, J.C., Tirole, J.: Two-sided markets: A progress report. *RAND J. Econ.* 37, 645–667 (2006).
 30. Gawer, A.: Bridging differing perspectives on technological platforms: Toward an integrative framework. *Res. Policy*. 43, 1239–1249 (2014).
 31. Henfridsson, O., Bygstad, B.: The Generative Mechanisms of Digital Infrastructure Evolution. *MIS Q. Manag. Inf. Syst.* 37, 907–931 (2013).
 32. Mocanu, E., Nguyen, P.H., Gibescu, M., Kling, W.L.: Deep learning for estimating building energy consumption. *Sustain. Energy, Grids Networks*. 6, 91–99 (2016).

33. Yun, S., Park, J.H., Kim, W.T.: Data-centric middleware based digital twin platform for dependable cyber-physical systems. *Int. Conf. Ubiquitous Futur. Networks, ICUFN*. 922–926 (2017).
34. Evans, P.C., Gawer, A.: *The Rise of the Platform Enterprise A Global Survey*. , New York (2016).
35. Gawer, A., Cusumano, M.A.: Industry platforms and ecosystem innovation. *J. Prod. Innov. Manag.* 31, 417–433 (2014).
36. de Reuver, M., Sørensen, C., Basole, R.C.: The digital platform: a research agenda. *J. Inf. Technol.* 33, 124–135 (2018).
37. Rajratnakharat, Bavane, V., Jadhao, S., Marode, R.: Digital Twin: Manufacturing Excellence Through Virtual Factory Replication. *Nc-Race* 18. 6–15 (2014).
38. Chesbrough, H., Vanhaverbeke, W.: Open Innovation : A New Paradigm for Understanding Industrial Innovation. *Open Innov.* 4, 1–27 (2005).
39. Chesbrough, H., Vanhaverbeke, W., West, J.: *New Frontiers in Open Innovation*. Oxford University Press, Oxford (2014).
40. Peltz, C.: *Web Services Orchestration and Choreography*. (2003).
41. Schulte, S., Schuller, D., Steinmetz, R., Abels, S.: Plug-and-play virtual factories. *IEEE Internet Comput.* 16, 78–82 (2012).
42. Liu, Z., Meyendorf, N., Mrad, N.: The role of data fusion in predictive maintenance using digital twin. *AIP Conf. Proc.* 1949, (2018).
43. Borodulin, K.: Towards Digital Twins Cloud Platform: Microservices and Computational Workflows to Rule a Smart Factory. 209–210 (2014).
44. Tao, F., Zhang, H., Liu, A., Nee, A.Y.C.: Digital Twin in Industry: State-of-the-Art. *IEEE Trans. Ind. Informatics.* 15, 2405–2415 (2019).
45. Mourtzis, D., Doukas, M., Bernidaki, D.: Simulation in manufacturing: Review and challenges. *Procedia CIRP.* 25, 213–229 (2014).
46. Müller, J.M.: Antecedents to Digital Platform Usage in Industry 4.0 by Established Manufacturers. *Sustainability.* 11, 1121 (2019).
47. Landolfi, G., Barni, A., Menato, S., Cavadini, F.A., Rovere, D., Dal Maso, G.: Design of a multi-sided platform supporting CPS deployment in the automation market. In: *2018 IEEE Industrial Cyber-Physical Systems (ICPS)*. pp. 684–689 (2018).