Information and Software Technology 169 (2024) 107424

Contents lists available at ScienceDirect

Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof

Stakeholders collaborations, challenges and emerging concepts in digital twin ecosystems

Nirnaya Tripathi *, Heidi Hietala, Yueqiang Xu, Reshani Liyanage

M3S Research Unit, University of Oulu, Oulu 90014, Finland

ARTICLE INFO

Keywords:
Digital twin
Digital twin ecosystem
Stakeholders
Systematic literature review
Empirical study
Definition
Software development

ABSTRACT

Context: Digital twin (DT) ecosystems are rapidly evolving, connecting many stakeholders, such as manufacturers, customers, and application platform providers. These ecosystems require collaboration and interaction between diverse actors to create value. This study delves into the collaboration of such stakeholders within DT-focused ecosystems.

Objective: This research aims to understand stakeholder collaboration within DT ecosystems, identify potential challenges, and provide insights for managing these stakeholders. It also seeks to define the DT ecosystem and its implications for both research and practice.

Method: A systematic literature review was conducted, supplemented by empirical evidence gathered from interviews with DT experts who were knowledgeable about the DT ecosystem. The study also analyzed DT systems, stakeholder roles, and the challenges with ecosystem-focused DT development.

Results: The study identified various stakeholders and their roles in adding value to a DT ecosystem. It highlighted the benefits of stakeholder collaboration, such as knowledge gain during DT system development. The research also revealed the technical and non-technical challenges encountered in ecosystem-focused DTs, emphasizing the importance of standardization as a solution. A new definition of the DT ecosystem was proposed, emphasizing its data-driven nature, interconnected DTs, stakeholder value creation, and technology enablement.

Conclusion: Stakeholder collaboration is pivotal in DT ecosystems, with each actor playing a distinct role. Addressing challenges, especially through standardization (OPC UA and ISO 23247), can lead to more efficient and coherent DT ecosystems. The insights provided by this study can guide industries in designing, developing, and maintaining their DT ecosystems, ensuring value creation and stakeholder satisfaction. Future research avenues that emphasize the importance of understanding the challenges involved and deploy appropriate solutions were suggested.

1. Introduction

In today's highly competitive industrial environment, Industry 4.0 helps provide a competitive edge to organizations by improving products and processes through the adoption of new technologies in industries. The benefits include cost minimization, product quality enhancement, scalability, and flexibility in production facilities. A digital twin (DT) system has become a promising and emerging area of focus in the Industry 4.0 era, through which industries can obtain a number of benefits, such as the ability to foresee problems in the development process and give early warnings, develop novel prospects, and design enhanced devices/products through digital representations [1,2]. The above benefits are realized by the physical and digital components, as shown in Fig. 1. The physical part contains actuators and sensors that collect data from physical objects (POs) in the specific industry sector, such as automotive and manufacturing, from which the data are transmitted via a wireless network and stored in data storage facilities. The data are accessed by the digital part to build a digital representation of the PO and to perform operations, such as simulations and analytics.

Another critical advantage of DTs is that they can integrate information from multiple sources and scales in real-time from the physical entities and create living models of these entities, which can be used for predictive maintenance [3]. In this aspect, the application of DTs can be seen in a number of areas (see Fig. 1), such as manufacturing, healthcare, industrial internet of things (IoT) environments, automobiles, retail, and smart cities. As computational power has improved, and the price of sensors has decreased, a number of emerging technologies, such as artificial intelligence, deep learning, and machine learning, as well

* Corresponding author.
E-mail address: nirnaya.tripathi@oulu.fi (N. Tripathi).

https://doi.org/10.1016/j.infsof.2024.107424
Received 1 September 2023; Received in revised form 12 February 2024; Accepted 12 February 2024
Available online 14 February 2024

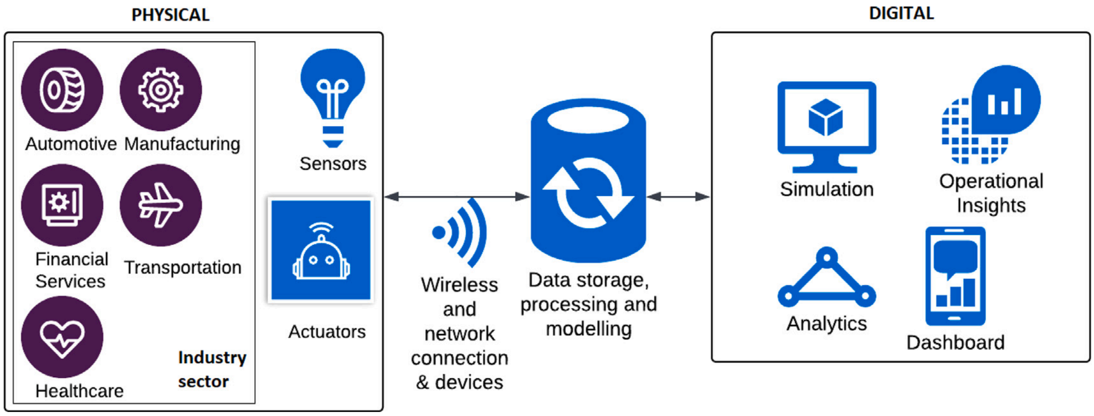

Fig. 1. Digital twin system.

as other trends, such as IoT and big data, have played an important role in the development of DT applications in the aforementioned areas [1]. Additionally, open-source frameworks, standards, and tools, such as Eclipse Hono (for IoT connectivity), Eclipse Vorto (for compiling and managing abstract device descriptions), Eclipse Ditto (for creating and managing DTs in the IoT), Unity (for creating interactive, real-time content), and Open Platform Communications Unified Architecture (OPC UA; data exchange standards for industrial communications), have made it possible and affordable to create DTs of POs [4].

1.1. Research motivation

Given the enormous business potential of DT systems to improve the efficiency and intelligence of manufacturing systems in Industry 4.0, DT research has focused on topics such as application contexts, life cycles, functions, architecture, and components/technologies [5]. However, an equally important topic for DT ecosystems is the complex and challenging process of developing a DT. Thus, interested stakeholders must be willing to integrate and cooperate in ecosystems that align with their missions. In this regard, a study [6] proposed the concept of a DT ecosystem that would boost the improvement of product and service development processes. This concept would direct organizations in identifying and developing novel product–service systems (PSSs). The authors further stated that the DT ecosystem would help in product design and life cycle management by creating value through an ecosystem of twin-driven products and services.

A benefit of working in ecosystems is that data from IoT sensors allow platform providers to develop DT systems. Once these DT platforms are established, application providers can build their DT services and applications around them for DT users. DT software systems and physical assets can potentially act as revenue generators. Services on top of DT systems can also yield more customers and users, allowing all contributing stakeholders to share revenue. Notably, realizing the potential revenue calls for understanding the various stakeholders involved.

In our earlier study [7], after a comprehensive analysis and synthesis of empirical data, stakeholders (see Fig. 2) and their requirements and roles in the DT ecosystem were identified and described. This DT ecosystem as a system of stakeholders denotes a DT environment that generates value for various related stakeholders, which include 13 potential stakeholders categorized as primary (e.g., manufacturers and subcontractors), secondary (e.g., maintenance service providers and platform integration service providers), and tertiary (e.g., research organizations and third-party value-added service providers). Some of them are described in the following:

- Suppliers deliver raw materials to manufacturers and utilize DTs to share relevant information about the supplies.

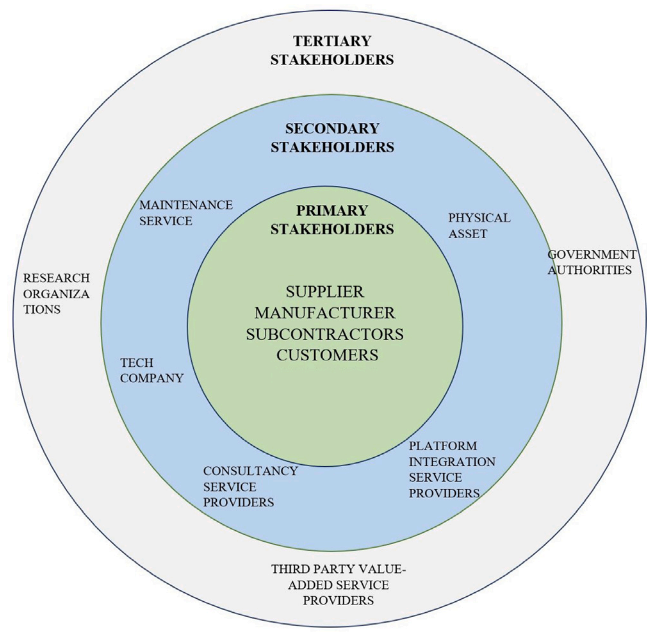

Fig. 2. Potential stakeholders within the DT ecosystem.
Source: Adapted from [7].

- Subcontractors perform tasks delegated to them by partners, including manufacturers, and utilize DTs to communicate information with these affiliated parties.
- Platform integration service providers facilitate the seamless integration of various DT systems and equipment from diverse stakeholders within the DT ecosystem. They enable these entities to collaborate, share data, and operate in harmony.

This study also presents the different stakeholders' requirements in detail. While our previous work provided insights into such a topic, DT ecosystems' stakeholder collaboration, challenges, and possible solutions remain under-researched, calling for a focus on empirical work and related literature.

1.2. Research objectives

Given the motivation outlined in the previous section, we put forward the following research objectives:

- To empirically study stakeholders' collaboration and the related benefits and challenges faced in DT ecosystems

- To explore the existing literature on DTs and DT ecosystems in order to gain state-of-the-art knowledge of the ecosystem context
- To understand the possible challenges and solutions at the ecosystem level

We conducted an empirical study and systematic literature review (SLR) to meet the research objectives. The study was carried out in the Operational eXcellence by Integrating Learned information into AcTionable Expertise (Oxilate) project to explore two industrial DT system developments (Digital Twin A and Digital Twin B), demonstrating that it is feasible to develop, implement, and deploy DTs on an industry ecosystem scale. Interviews were conducted with participants to address the research objective from practitioners' points of view. The driving factors for the ecosystem's functions in the project are as follows:

- Demand from customers and end users of the DT system, which is developed by project partners
- The development of DT standardization and technical frameworks
- The opportunity for joint value creation

To incorporate theoretical perspectives, this study uses a systematic approach to the DT ecosystem literature to explore the challenges involved and possible solutions. Accordingly, SLR guidelines, which were provided by Kitchenham and Charters [8], are followed.

The aim of this research is to add new knowledge to the literature on DT ecosystems by identifying stakeholder interactions, challenges, and possible solutions for these challenges in the context of DT ecosystems. Therefore, the results from the empirical study and SLR provide new theoretical knowledge while opening new avenues for further research in this area. We identify the following interesting concepts that emerge from the study, which we discuss in the end: (1) dynamic DTs (from adaptability to evolvability), (2) low-code development to increase the user base, and (3) data-centric development to articulate the business case and stakeholder benefits.

2. Background literature

2.1. Industry 4.0 and digital twins

Given their increasingly competitive nature and the need to recreate value in global industrial networks, Industry 4.0 has become a hype word in both academia and practice in recent years [9]. The term "Industry 4.0" was first introduced in 2011, referring to the Fourth Industrial Revolution in a broader context [10]. Its predecessor – Industry 3.0 – focused mainly on the automation of physical systems by reducing and replacing human interference with machines and programs to improve efficiency and performance. With the move from Industry 3.0 to 4.0, along with the digital transformation of industrial facilities, data orientation has become imperative. Here, more focus is given to the large amount of data generated in the industrial processes and communication between machines. Industry 4.0 also refers to processing these data to generate useful information that could be used in the industrial environment [10].

While the Industry 4.0 paradigm consists of a few characteristics, namely, decentralization, virtualization, interoperability, real-time capability, and modularity, cyber–physical systems have become significant elements in achieving virtualization. Moreover, as DTs have the ability to add value by facilitating the real-time monitoring capabilities of these real-world systems, they have a critical role in the context of Industry 4.0, allowing smart products and manufacturing systems to improve the competitive advantages of industries [2,4].

The initial DT concept was coined by Michael Grieves in 2002. He proposed a "conceptual ideal for product line management", a model that had the characteristics of DT, namely, real space, virtual space, the link for data flow from real space to virtual space, the link for information flow from virtual space to real space, and virtual subspaces [11]. Later, the term "digital twin" was put forward by NASA in its integrated technology roadmap [12].

2.2. Definition and classification

Digital twins (DT), as an emerging area of research, have many definitions in the literature. As seen in Table 1, DT definitions have evolved with varying emphases. Despite the differences in these definitions, they also have commonalities. In simple terms, a DT can be defined as a full digital representation (virtual counterpart) of a physical system or a product (physical counterpart) in which the virtual and physical counterparts are connected with each other in real time to transfer data from the physical counterpart to the digital one and vice versa.

2.2.1. Classification

The diverse DT definitions presented in the previous section also provide opportunities to develop different types of DTs based on such definitions. Depending on the level of integration of these virtual models in the system, they can be categorized into *digital models, digital shadows, and DTs* [14].

A digital model is a virtual representation that does not have any type of data transfer between it and the physical system. In virtual representations, the data transfer is manual, so changes in the state of the physical system are not dynamically represented by the virtual counterpart [15]. Its exemplary usage can be simulation models of designed future factories, mathematical models of new products, or any other virtual models of a physical counterpart that do not use any form of automatic data transfer [14].

A digital shadow comprises an automated one-way data transfer between a PO and its digital counterpart. In this virtual representation, a change in the physical system's state is dynamically represented in the digital object. However, changes in the state of the virtual representation are not dynamically represented in the physical system, as data transfer is only one way. [14].

Following this classification further, a DT can be seen as a digital representation of a PO that facilitates automated bidirectional data exchange between the physical system and its virtual counterpart in real time. In this type of digital representation, the virtual counterpart controls the PO and thus affects it. A change in either system dynamically affects the counterpart and vice versa [15].

A further classification of DTs was given by Kritzinger et al. [14], who described them based on their intended uses in the production/process life cycle. Accordingly, DTs can be classified as product or predictive twins, product or production twins, or performance twins. Based on this classification, predictive twins comprise all the artifacts of a product used mainly for predictive tasks, while predictive twins comprise the models for the digital process or the system. Performance twins are focused on improving the performance of processes/systems by analyzing the actual performance and utilization data [12].

DTs can also be classified as digital twin prototypes (DTPs) and digital twin instances (DTIs), as defined by Grieves and Vickers, 2017 [11]. A DTP is a prototypical physical artifact that consists of information sets capable of describing the PO. These sets of information can be, for example, requirements, 3D models, or bills of material of the PO. On the other hand, a DTI is described as a "specific corresponding physical product that an individual Digital Twin remains linked to throughout the life of that physical product".

While the above studies provide various viewpoints for classifying DTs, the scope of the twin requires attention in order to study DTs holistically. A recent study [15] put forward three organizational scopes for DTs: internal operations, value chain, and ecosystem. While the first accounts for a DT as internal to the focal organization, the latter two extend the scope to the network and ecosystem levels accordingly. Consequently, the authors highlighted the importance of the tight coupling of DTs as software platforms with the service development ecosystems. This is critical for promoting and facilitating the widespread use of DT applications in the Industry 4.0 era.

2.3. Digital twin ecosystem

Initially, the ecosystem concept was developed in biological research before it spread into many other significant areas, such as

Table 1
Digital twin definitions.

Study	Definition	Explanation
[12]	*"[a] very realistic model of the current state of the process and their own behavior in interaction with their environment in the real world"*	The study states that DTs are capable of predicting product behavior aside from producing a mere representation of it and that they provide room for optimizing products and processes as a result of the speedy simulation of vast amounts of data in real time.
[11]	*"a set of virtual information constructs that fully describes a potential or actual physical manufactured product from the micro atomic level to the macro geometrical level"*	The study states that the virtual information construct is connected with the physical system during the entire lifetime of the system. According to the authors, the DT provides any user with the ability to obtain information from the physical system when the DT is at its optimum level.
[13]	*"a multi-physical, multi-scale, and probabilistic simulation model of a complex product. It uses updated sensors and physical models to mirror physical life in the digital world and vice versa."*	The study mentions that for a virtual model to be considered as a DT, it needs to possess the characteristics of scalability, interoperability, and fidelity.

technology, economics, and management [16]. The ecosystem concept has evolved according to the context of its use and appeared in terms such as "business ecosystems", "software ecosystems", and "IoT ecosystems". These concepts emphasize networks of organizations driven by common objectives, such as creating high-quality products that benefit customers. In terms of business ecosystems, an ecosystem can be characterized as "a large number of loosely interconnected participants who depend on each other for their mutual effectiveness and survival" [17]. These ecosystems typically connect different organizations in product design, production, distribution, or implementation.

In exploring DT ecosystems, Silva et al. [6] defined a DT ecosystem as multiple interconnected instances of a DT or different DTs that have been arranged into value chains using different enabling technologies. The concept of value chain in DTs involves using DT technology in every step of product or service creation and selling, while the ecosystem is like a network of these DTs linked together, containing technology, processes, and stakeholders. These ecosystems, for example, in a manufacturing environment, can consist of different facilities, such as one or more entities that are responsible for different functions of sales, maintenance, and so on. The different parties compose the value chain and the ecosystem, including the data generated by each party [6]. A DT ecosystem can also be described as the environment that includes a single DT, sensors, the technologies used, users, and other components [3].

One of the fundamental concepts related to DT ecosystems is the development of a real twin capable of connecting the physical system with its virtual mirror. This concept could potentially lead industries to achieve real-time prediction and repeated and continuous optimization of the different parameters in a system within the ecosystem. It would further enable risk warnings in advance, fault detection, and intelligent optimization instructions for various categories of workers, such as system operators and maintenance workers [18].

In 2015, General Electric and its partners in aircraft manufacturing developed a DT ecosystem that facilitates a collection of DTs and different industrial services of the organizations related to the engine, airframe, and further systems, which, in turn, created a more comprehensive and advanced DT of the complete aircraft [1]. According to the study, the DTs communicated continuously, so they learned from each other continuously. The main objective of developing this DT ecosystem was to enable early cautions and failure forecasts for different major components of the aircraft. Accordingly, it provided the power of accurate prediction of the future [1].

In 2019, Rosen et al. [19] demonstrated how DT development, as well as the following transfer to partners and stakeholders in various value chains, can create an ecosystem while the collaboration and exchange of digital artifacts dynamically generate mutual benefits. The study provided a technological vision of a DT ecosystem and its application in the field of mechatronic systems. The authors presented the characteristics of a digital ecosystem in which organizations and people share digital artifacts and models using electronic platforms. These DT-based digital platforms can support product design and life cycle management while simultaneously generating value through an ecosystem of twin-driven PSSs [6]. As suggested by Silva et al. [6], these platforms must have data layer and data orchestration tools, which are crucial for integrating the data recovered from the multiple instances of DTs and for fully realizing the interoperability between different twin systems.

Another study on DT ecosystems was conducted by Pantelidakis et al. [20], in which they focused on creating and applying a new DT ecosystem within additive manufacturing. It emphasized the use of a real-time development platform to improve manufacturing procedures and to demonstrate practical use cases and results.

A potential DT ecosystem model (Fig. 3) can be described as follows [21]:

- A PO is a physical representation of a product that is sold or leased to customers. With a contract, the PO provider or customer can allow the digital platform provider to access the data, communication, storage, and processing capabilities needed to support the DT concept.
- The digital platform provider can instantiate the PO's (i.e., the logical object) software version. The interfaces and views on logical objects can be used to create services. The combination of the physical and logical objects comprises the DT system.
- The digital platform provider can grant an application provider access to the DT to build services that will be provided to end users.

3. Methodology

To achieve our research goals, we conducted an empirical study and SLR focusing on stakeholder interactions in DT ecosystems. The research process is presented in Fig. 4. In the following section, we present the research questions guiding our interviews and SLR, followed by descriptions of both the research processes and the data analysis.

3.1. Research questions

- RQ1. How do stakeholders collaborate and benefit in digital twin ecosystems for digital twin development?
- RQ2. In digital twin ecosystems, what are the main challenges faced during digital twin development?

Table 2
Interviewees details.

ID	Location	Focus	Software experience (years)	Role description in relation to DT development
I1	Belgium	General	1–2	The interviewee is a machine learning and NLP solutions research engineer.
I2	Belgium	General	5–10	The interviewee works as a research manager to develop DT simulation.
I3	Turkey	DT A	5–10	The interviewee has been working on DT projects in the automotive and food sectors.
I4	Turkey	DT A	1–5	The interviewee is a software engineer working to develop DTs in the automotive sector.
I5	Turkey	DT A	5–10	The interviewee is a consultant for creating DT software for an automotive company.
I6	Finland	DT B	10–20	The interviewee is a research director who aims to build DT prototypes.
I7	Finland	DT B	5–10	The interviewee provides solutions based on data from DTs to optimize the process.
I8	Finland	DT B	5–10	The interviewee has experience in developing DT systems for testing autonomous vehicles.

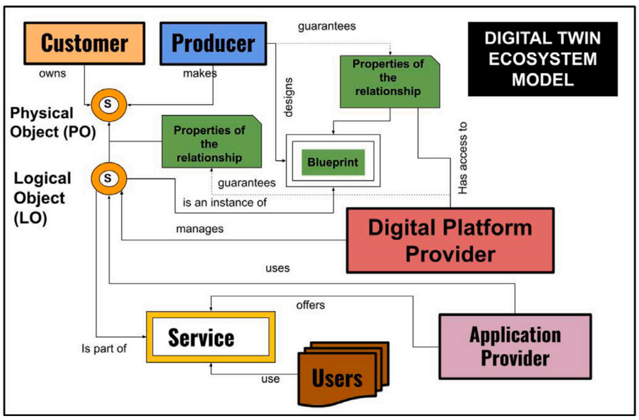

Fig. 3. Potential DT ecosystem model.
Source: Adapted from [21].

3.2. Empirical study

To meet our research objectives, we conducted an empirical study in the Oxilate project. Empirical evidence was collected to obtain an in-depth understanding of the DT ecosystem phenomenon under study by exploring two industrial DT system development study units (Digital Twin A and Digital Twin B), demonstrating that it is feasible to develop, implement, and deploy DTs on an industry ecosystem scale.

- Digital Twin A: Four companies worked jointly for DT development in a DT ecosystem for over six months. The organizations are located in Turkey. These organizations conducted DT development in an ecosystem environment, so they were suitable for providing information on the research objectives.
- Digital Twin B: The second unit is located in Finland. As a large organization developing a DT system for six months, the unit was considered appropriate for providing data for our research objectives.

3.2.1. Data collection

Empirical data were collected through interviews. Interviewees were selected based on their knowledge of and experience in DT development. They were working in the case companies involved in the research project that focused on DT development. Thematic, semi-structured interviews were conducted with systems engineers, senior software designers, and project managers to study DT development. These interviews were held online in English and were audio-recorded. An interview guide was designed for the data collection. As the interviews proceeded, our interview guide evolved to gain maximum coverage and depth for our study. Table 2 summarizes the details of the interviewees, and Appendix B presents the interview questions.

3.3. Systematic literature review

We used a systematic approach to the DT ecosystem literature to incorporate theoretical perspectives into the empirical findings. As per Kitchenham and Charters [8], there are several reasons for performing an SLR, such as to summarize the available literature on a topic, to determine the current research gaps on a certain topic, or to develop the foundation for new research in that area. We chose the SLR approach to study the existing literature and to understand current knowledge on DT ecosystems. We used the Parsifal SLR tool to conduct our literature review.

3.3.1. Planning

The review protocol used in this study was developed in the planning phase. When finalizing the review protocol, attention was given to the aspects of the protocol, such as Keywords and Inclusion Selection Criteria. The search string structure according to Kitchenham and Charters [8] in terms of population and intervention was specified below.

- Population: "Digital twin" AND (Challenges OR Issues)
- Intervention: Ecosystem

Electronic databases, such as Scopus, IEEE Xplore, ACM Digital Library, and Science Direct, were selected for the study. They were chosen because they are widely used in the software engineering domain and allow for advanced query searches. The inclusion Selection Criteria of papers were as follows.

- The publication language is English.
- The publication is part of any journal, conference, or workshop.
- The words "digital twin" and "ecosystem" appear in the abstract, title, and keywords.
- The full text is available for reading, and it discusses digital twin ecosystems.

3.3.2. Conducting and reporting

In the initial search of the databases, 497 studies were found. After 181 duplicate studies were discarded, 315 were left for further assessment. Applying inclusion and exclusion criteria to each paper's title and abstract diminished the pool of papers for full-text reading to 237. Finally, 10 papers were selected as primary studies after the exclusion and inclusion criteria were used for quality assessment. Fig. 5 gives an overview of articles retrieved and selected from the four database sources.

3.3.3. Quality assessment

The quality assessment criteria used in this study were developed based on the guidelines provided by [22,23], which were as follows:

1. Rigor
2. Relevance
3. Reporting

In the selection of studies based on the quality assessment criteria, it is expected that a study will have an overall score above 8. As such, quality assessment criteria were applied to the set of 10 primary studies. Table 3 presents the list of primary studies included in the data extraction process, along with their quality assessment scores. Primary articles description is provided in Appendix A.

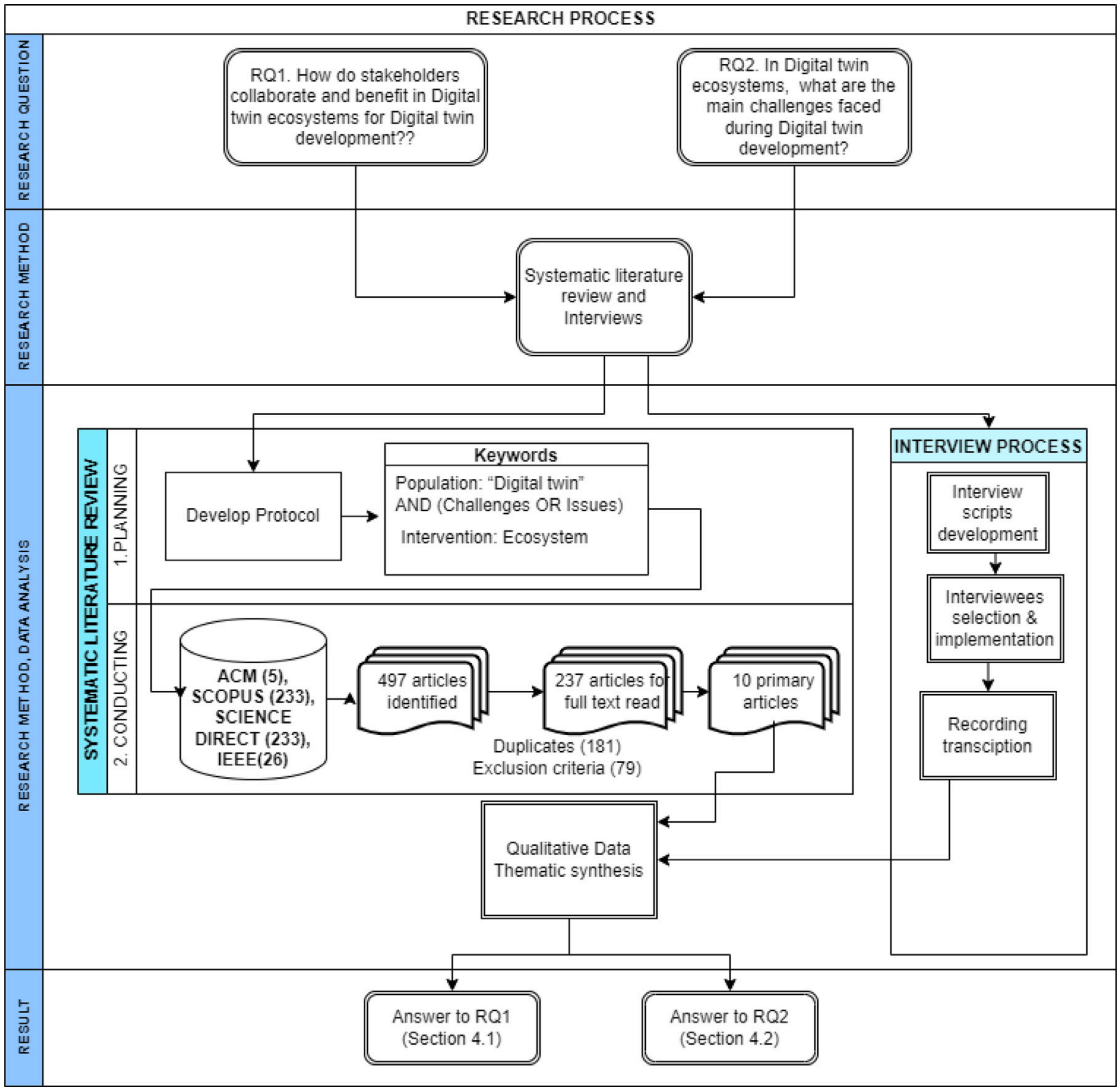

Fig. 4. Research process.

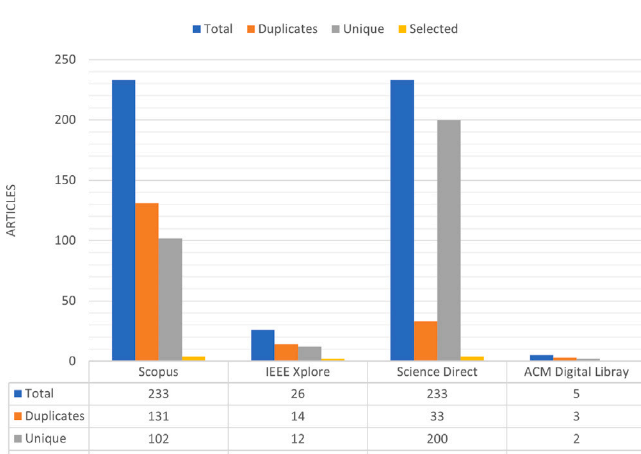

Fig. 5. Overview of articles search and selection per database sources.

Table 3
Primary studies overview.

ID	Source	Year	QA	Database
P1	[24]	2019	7	Scopus
P2	[25]	2021	7	Scopus
P3	[26]	2019	8	IEEE Xplore
P4	[27]	2021	9	ScienceDirect
P5	[28]	2021	8	Scopus
P6	[29]	2021	9.5	ScienceDirect
P7	[19]	2019	7	ScienceDirect
P8	[6]	2021	8	IEEE Xplore
P9	[30]	2021	7.5	ScienceDirect
P10	[15]	2020	7.5	Scopus

3.4. Data analysis

We conducted data analysis in two stages. We used the deductive thematic approach using the guidelines by [31]. Data analysis was performed on qualitative data generated from the interview transcripts and primary articles to answer our research questions and achieve the goals of our study.

Table 4
Extracted data classification.

Data	Field	Addressed in paper
Article title, Description	String	Appendix A
Publication year, Database source	Date, string	Fig. 5
Interviewee details	String	Table 3
Digital twin discussed	String	Section 4.1.1, 4.1.2
Stakeholders interactions	String	Section 4.1, RQ1
Digital twin ecosystems challenges	String	Section 4.2, RQ2
Guideline suggestion for challenges	String	Table 7

3.4.1. Stage 1

In the first stage, interview audio file transcriptions were done using a professional transcription company. The authors reviewed the transcription by listening to the audio files to fill in missing words and correct unclear sentences in the transcriptions. Similarly, the full-text files of the primary studies were downloaded to conduct the data analysis. As shown in Table 4, a data extraction classification was used when collecting data from the primary studies and interview transcripts. After this, the data were extracted into an Excel sheet for deductive analysis in the second stage.

3.4.2. Stage 2

The authors followed a deductive coding approach, creating codes to seek relevant information from the extracted data. These codes focused on DT units, stakeholder interactions, and challenges. The authors reviewed the extracted data from the primary studies and interview transcripts several times to identify all relevant code texts. The aim was to structure the extracted data to be adopted in reporting the state of the art and empirical knowledge in the field in order to attain our study objectives. The findings and interpretations are provided in the Results section.

4. Results

This section focuses on the collaborative efforts of stakeholders (4.1), specifically highlighting their roles and functionalities considering Digital Twin A and Digital Twin B. Following this, Section 4.2 addresses the inherent challenges in a DT ecosystem.

4.1. Stakeholder collaborations

4.1.1. Digital Twin A

This DT system focuses on minimizing the energy (electricity and natural gas) and chemicals consumed in the vehicle painting process. The stakeholders involved in developing the system, as shown in Fig. 6, are Org A (customer), Org B (platform provider), Org C (technical partner), and Org D (technical partner). The system uses integrated data collected from various sensors and robotic arms, thus enabling users to track the progress of activities in the painting process.

Stakeholder collaborations: Org A's vision was to use existing sensors and implement new ones in the paint shop of one of its manufacturing plants to provide an experimental foundation for data collection from a real-time machine park. Multiple parameters from various machines must be collected from SCADA systems and directly from various sensors and then transferred to a central server and database. The database provides data for three software modules: machine learning (ML), DT, and stochastic analysis. The central database (Postgre SQL and Mongo DB) architecture was constructed in a joint project with Orgs B, C, and D.

Each software module delivers specialized reports for the software products to be developed: (1) Digital Twin A, (2) simulation, and (3) dashboard. The reports facilitate the management of manufacturing plants, result in better decisions, and help take action with a virtual assistant. Org B uses the data to create and run Org A's paint shop process DT.

Collaboration in the ecosystem, as can be seen in Fig. 6, benefits stakeholders as follows:

- Org A runs the DT solution in its paint process plant.
- Org B has a new product output as a DT implementation framework for processes and dashboards.
- Org C gains expertise in ML applications for industrial problems.
- Org D develops new algorithms for natural language processing techniques and uses them for new products.

4.1.2. Digital Twin B

Digital Twin B's system was developed by Org X (platform and application provider) in Finland. It is intended to support fully integrated systems in customers' operational workflows, which is achieved by developing and integrating actionable data analytics with expert knowledge of manufacturing maintenance. While it provides support for professionals, it also creates direct business value in the product life cycle it serves. Digital Twin B's architecture and its components are shown in Fig. 7.

The three components are as follows:

- The user interface boasts several key features to enhance user interaction and content management. It provides a visually engaging experience by displaying 3D models alongside part names, facilitating better understanding and navigation.
- The web server, operating as a REST application programming interface (API), serves as a central hub for various functions. It effectively stores and updates information related to states, animation templates, and variables. Users can initiate animations and leverage predefined JavaScript functions, simplifying development.
- The adapter, combining the OPC UA Client and REST API functionality, acts as an intermediary between systems. It retrieves data from an OPC UA server and maps OPC UA variables to facilitate seamless data exchange and integration between these two systems.

Stakeholder collaborations: Based on [32], a DT reference model demonstrating how a DT, which is a logical object, and a PO communicate, we determined that humans are critical parts of this dynamic DT. Digital Twin B shifts from the dynamism of adaptability in a stand-alone DT system to the dynamism of evolvability in a DT ecosystem.

First, with human expert knowledge, Digital Twin B can adaptively detect the collisions of POs, such as pipes, before a plant is built and installed. The locations of safety measures can be made visual and be sanity checked using visualization technologies, such as virtual reality (VR) and augmented reality.

Second, metadata, real-time data, and other models are combined to realistically represent the plant and its control flows in operation. Such data not only allow Digital Twin B to improve its system and customers' operations adaptively but also enable Org X ecosystem partners and stakeholders to develop more solutions that work with Digital Twin B, thus creating both local system adaptability and global ecosystem evolvability.

Third, from a software development perspective, Org X uses low-code and no-code implementations, which means building the DT through pre-defined API and JavaScript functionality to enable DT evolvability that integrates ecosystem stakeholders and third-party client solutions. Furthermore, operational data are mapped back to the metadata produced by Digital Twin B's ecosystem partners, and provide knowledge-based data to the users of org X's MediaWiki as part of the DT. More specifically, Digital Twin B's architecture is designed to place a PO identifier in VR, and it contains a soft tag, which is actually the OPC UA Server nodeId (selected from a list in the OPC UA server).

4.1.3. Lessons learned

In the case of Digital Twin A, stakeholder collaboration was viewed as a vital part of the development process. While it was critical for the customer and technology partners working on the DT, involving other vendors, such as those for legacy systems, was often necessary.

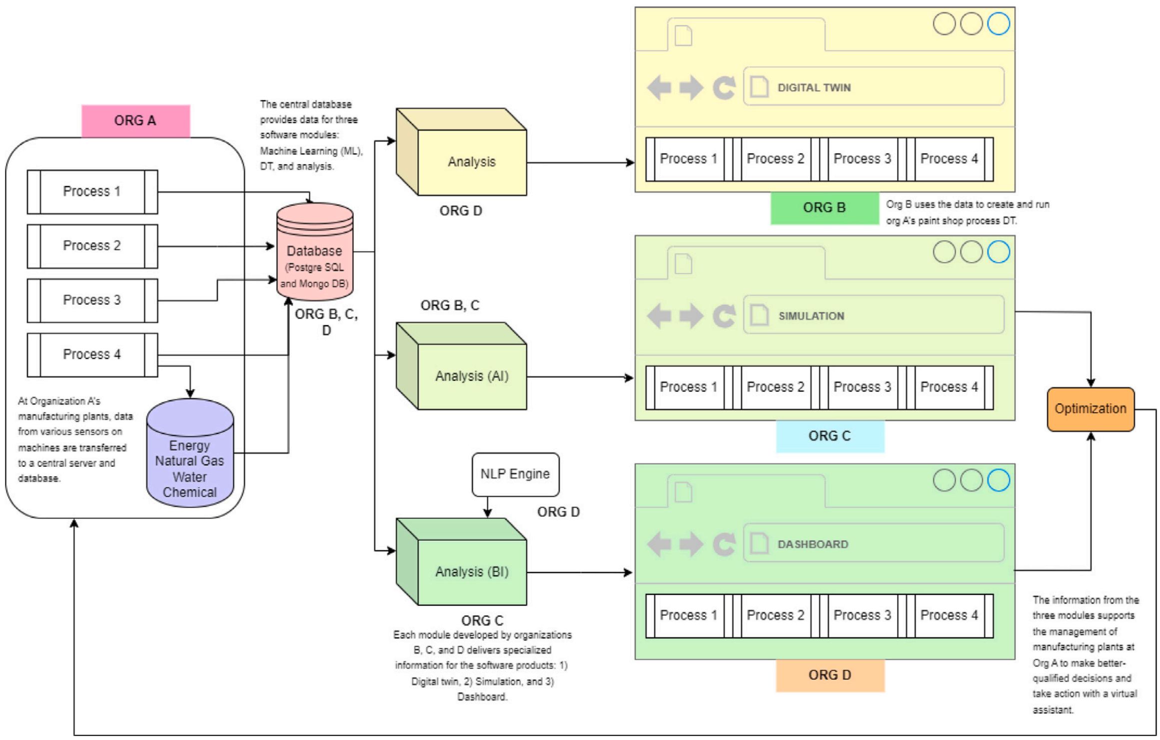

Fig. 6. Stakeholders' collaboration in Digital Twin A's development.

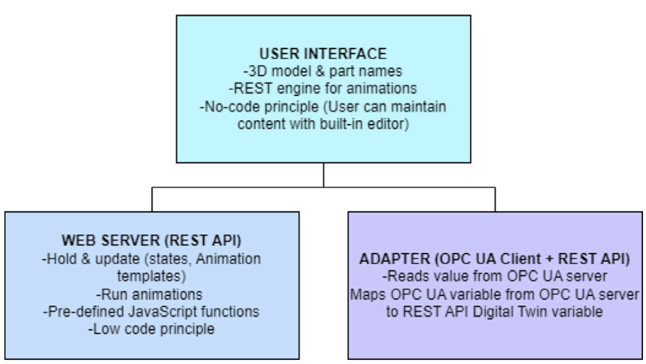

Fig. 7. Digital Twin B's architecture.

Adopting an ecosystem perspective was considered suitable for DT development to meet this need. The goal of partnering for DT development for a specific customer was eventually to offer solutions to other businesses. Thus, the primary aim was to foster an ecosystem-centric collaboration that serves a more extensive clientele and surpasses a particular value chain (cf. [15]).

In the case of Digital Twin B, collaborative development among stakeholders was considered beneficial, providing advantages for the partners involved. Stakeholder collaboration resulted in an improved understanding of customers' needs, challenges, and objectives, aiding partners in refining their products and services. However, while benefits were identified, the stakeholder collaboration concerning the DT appeared more like a supply chain than an ecosystem-centric software and service development (cf. [15]).

In summary, our empirical research showed a range of stakeholder collaborations in the organizational scope of DTs development. While the collaboration (Digital Twin B) reflected a more conventional customer–vendor dynamic, in which the customer integrates various services and technical solutions and dictates the partners contributing to the solution, the ecosystem approach (Digital Twin A) offers partners a pathway to grow their businesses and uncover new business prospects, as evidenced in our cases.

Table 5
Potential Challenges in the DT Ecosystem.

Challenge	Primary study source	Interviewee source
Interoperability	P2, P10	I3
Cybersecurity	P1, P3, P4, P9	I1, I3
Data governance and management	P1, P3, P6	I5, I6, I7, I8, I9
Data sharing and fusion	P2, P5, P9 P10	I7, I8
Other challenges		
Orchestration	P5, P10	–
Gaining stakeholder commitment	–	I1, I8
DT expertise	–	I5, I7,I8

4.2. Challenges in the DT ecosystem

Our data analysis from the interviews and SLR revealed several challenges that can be encountered in DT ecosystems. These challenges are described in Table 5.

4.2.1. Interoperability

Our empirical analysis revealed that the interoperability of different DT systems is critical. Interoperability, which refers to the ability of different DT systems to seamlessly integrate and communicate despite using different standards or protocols, poses challenges, such as harmonizing diverse technologies and standards across various companies within the DT ecosystem. Thus, seamlessly connecting the systems used by different ecosystem partners is often challenging. While DTs rely on using different data sources and technologies, challenges arise, particularly when they need to be integrated into different systems, devices, and data sources using different standards or protocols. This opinion was also reported by Interviewee 3:

"One of the most important things is making these digital twins helpful for different companies. In the ecosystem, different digital twins will have to integrate to communicate with each other, and it will be a challenge to create some regulations or standards for this, and [it] will take a long time".

The literature also considers integrating and interoperating systems to be major difficulties in DT ecosystems. The technologies used are expensive and complex, complicating the use of DTs. Most industrial software applications are proprietary, preventing industry players from embracing open-source software [P10]. Providing interoperable solutions and platform agnostic is essential to serve stakeholders from various disciplines and domains [P2].

4.2.2. Cyber security

Cyber security pertains to the protection of DT systems from cyber threats. Because of the diverse interactions within DT ecosystems, they can be vulnerable to cyber-attacks, such as hacking and phishing. Thus, cyber security involves such activities as safeguarding data storage, access, sharing, and authenticity within these complex systems.

Given the complexity of DT-based systems, that is, physical-to-virtual, physical-to-physical, virtual-to-virtual, and virtual-to-physical interactions, and enormous product and service data, managing them efficiently and securely from the perspectives of data storage, data access, data sharing, and data authenticity is challenging. Security is one of the major challenges in DT ecosystems, making them potential subjects to cyber-attacks [P3].

Ensuring communication security among the components of DT ecosystems is also critical, as identified by [P1], who proposed a DT ecosystem design for the healthcare industry. The authors highlighted that measures need to be applied at the network level to prevent attacks during data transmission between the main DT components. Similarly, another study [P3], which proposed a DT ecosystem model, identified that the DT ecosystem is vulnerable to cyber-attacks, such as hacking and phishing. Accordingly, the credibility of the generated data has to be secured, and any cyber-attack possibilities should be diagnosed early on and prevented. Therefore, while security is a challenge in DT ecosystems [P9], security and credit issues in the system are also critical in enabling the sustainability of businesses [P4].

In summary, cyber security is a critical issue because DTs rely upon different data sources, networks, and communication protocols to perform, and this, in turn, can create vulnerabilities. For example, data breaches related to POs or to users and stakeholders, and unauthorized access to confidential information impose risks for DTs.

4.2.3. Data governance and management

In DT ecosystems, data governance and management encompass such aspects as acknowledging proprietary rights, managing stakeholder access and trust, addressing confidentiality, adopting decentralization, customizing data privacy, overcoming infrastructure challenges, and utilizing data for insights and enhancement. [P3] discussed in their research that health data pertaining to any asset's systems (in this context, an aircraft) are proprietary to suppliers. Distributing this information without infringing upon intellectual property rights poses a complex challenge. This limitation can prevent the DT from fully representing its physical counterpart. Therefore, data governance is a challenge in DT ecosystems.

The relevant stakeholders involved in the data life cycle need access to the DT in order to manage DT data. Although the stakeholders participating in these processes work together, they pursue different goals. Trust when sharing data via the DT is not given by default. As a result, confidentiality and access control issues arise in DT ecosystems. To address these problems effectively, a decentralized approach may be needed, particularly in large-scale environments with multiple tenants [P6]. In an eHealth context, [P1] proposed that stakeholders should be aware of the information that DTs disclose, which necessitates having highly customizable and intelligent privacy profiles. Furthermore, the study reported that DTs should be capable of determining the information that will be shared with other DTs and/or smart services, as different information will have varying sensitivity levels.

Accordingly, data governance and management and the proper use of data are challenges encountered at the DT ecosystem level because they may introduce a more extensive range of ecosystem stakeholders with larger amounts of data within the DT ecosystem. As DT uses different data sources from sensors, devices, and different systems, data need to be stored and analyzed in real time for DT operation. Some issues could be related to data integration because of data arriving from different sources; data quality in terms of accuracy, completeness, consistency, and data storage; and processing infrastructure. In relation to this, Interviewee 5 made the following comment:

"I'm talking about these kinds of smart factories, but it won't happen unless we collect data properly if you can manage the data, provide the insights, and give the best out of that data and use it to achieve better production in the next round; even this part of the story is quite hard to implement all over the world".

4.2.4. Data sharing/data fusion

Data sharing, in the context of DT ecosystems, involves sharing data, whereas data fusion involves combining the shared data from various sources and stakeholders for improved decision making and product development. The identified challenges related to these concepts include data silos, access, integration, ownership, and protection of sensitive information. Data sharing can be classified into internal and external data sharing. While data sharing across the value chain can bring tangible benefits and transparency to ecosystem partners, it has proven to be challenging because of silo effects. While data sharing can be a critical difficulty that is beyond DTs' technology and engineering complexities, data access can be equally problematic in pursuing ecosystem-focused DTs [P10]. Thus, data access related to data provisioning, utilization, and sharing was identified as another key challenge in DT ecosystems.

The data fusion of DTs involves three processes: data preprocessing, data mining, and data optimization [P10]. Accordingly, the lack of integration capability in these diverse parameters gives rise to data ownership, cleansing, and fusion challenges. The heterogeneity of the DT system may also include the involvement of third-party stakeholders, adding to the complexity of designing an all-encompassing DT. Notably, the ultimate goal of involved organizations is to leverage the available data to build faster, more cost-effective, and high-quality products [P5]. However, with the successful utilization of their data, automotive manufacturers, for example, may have different maturity rates. Accordingly, accurate acquisition and rigorous data logging can be challenges to implementing a sophisticated DT ecosystem.

The process of data fusion within a diverse and dispersed ecosystem can present significant challenges. While various stakeholders may be cognizant of one another's data, the understanding and interpretation of information linked to the same real-world object can significantly differ. This discrepancy complicates data integration and may create issues when connecting data across multiple partners [P2].

Data sharing was also identified as a challenge in our empirical data on DT ecosystems. The informants raised concerns regarding competitors using the same API. Interviewee 7 commented as follows:

"What happens if there are competitors using the same API? That will perhaps be a challenge because they want to access almost the same data, and then there will be sensitive data for those partners".

Because of the nature of the DT ecosystem as a wider platform, the possibility of data being transmitted to unwanted users can arise. Notably, deciding which data should be opened to which partner would be problematic because the same partner may also have some stake in the data from other sources.

4.2.5. Other challenges

- Orchestration

  Based on the findings of the SLR, orchestration in DT ecosystems refers to coordinated interactions among users, companies, physical entities, and virtual models to solve complex problems. According to [P5], openness and governance toward orchestrating collaboration among different users and companies are required in DT ecosystems. DT components must interact and work together to solve complex problems. The interaction and collaboration can be of three types: 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 any single entity cannot perform. Multiple virtual models can be connected with a network for information sharing through virtual-to-virtual interaction. Using virtual-to-physical interaction and collaboration, the virtual model can be tuned in synchronization with the PO, while the PO can be dynamically adjusted based on direct commands from the virtual model [P10]. These interactions need to be orchestrated to maintain openness and governance. Thus, orchestration poses a challenge in DT ecosystems.

- Gaining stakeholder commitment

  Stakeholder commitment refers to the willingness of individuals or companies to support and invest in DT development in an ecosystem, which involves dedication regardless of the potential challenges involved, such as justifying the need, securing funding, and dealing with uncertainties about project outcomes and added value. One of the main non-technical challenges, according to Interviewee 1, is justifying and explaining to stakeholders the need to work in an ecosystem. The interviewee also highlighted that it would be difficult for companies to justify the investment in this kind of project without knowing its actual output. Accordingly, I1 stated the following:

  "*The biggest challenge would be to explain to various stakeholders why this DT ecosystem is needed. That's the first. Second, I think it would be quite problematic to find the funds because I don't think there would be a lot of companies or a lot of, again, stakeholders who would be ready to pay for this kind of solution without really knowing what the output is*".

  Similarly, Interviewee 8 highlighted the same challenges as Interviewee 1 did; that is, it is difficult to increase the understanding of non-technical experts, such as business managers, and see why this move toward the DT ecosystem would be beneficial. Specifically, the following was stated:

  "*Justifying the usability and added value of DT for non-experts would be a challenge.*"

  Many interviewees recognized the benefits of working in an ecosystem-level collaboration regarding DT development and operation. However, many stakeholders perceived the move as cumbersome, for example, in terms of the needed investments, unclear project outcomes, and potentially limited added value. These issues were observed to hinder stakeholders' interests and commitment, challenging ecosystem-level DT development.

- Digital twin expertise

  DT expertise, based on our empirical evidence, refers to the specialized knowledge and skills necessary for developing and advancing DTs in ecosystems. Finding skilled people to develop these complex systems that DTs represent was identified as another challenge by the interviewees. In this context, Interviewee 5 stated the following:

  "*We'll need intelligent people; that's going to be hard to find. Intelligent people are scarce resources.* "

  Interviewee 7 also highlighted the need for expert or skilled human resources when developing these DT ecosystems. Skilled human resources are needed during the different phases and in the different areas of the DT ecosystem. Interviewee 7 commented on the matter as follows:

  "*We must understand how the process works before we can do the virtual counterpart or the ecosystem of digital twins. And that requires, usually requires, special expertise and knowledge, and it can be that only a few people within a company really understand how the process works.*"

  Based on our empirical data, understanding DT-related processes and physical DT counterparts plays a critical role in DT development. A high level of expertise is needed in organizations, but simultaneously, finding skilled people may be a challenge that hinders the development process toward ecosystem-level DTs.

5. Discussion

5.1. Stakeholder collaboration and benefits (answer to RQ1)

DT ecosystems connect stakeholders, such as manufacturers, customers, suppliers, sub-contractors, external maintenance service providers, and intelligent robots, calling for collaboration and interaction between these diverse actors. In this context, our study has revealed that these stakeholders have different roles and unique requirements to be considered for value creation in DT-focused ecosystems, as identified by the literature as well [7].

In Digital Twin A, an ecosystem-focused DT as defined by [15], joined together different stakeholders to develop a DT system for a manufacturing company. One benefit is that the participants who create the DT system can enhance their understanding of the system's development and increase the potential revenue beyond a specific value chain. By contrast, Digital Twin B resembled more a value chain network transitioning toward an ecosystem, showcasing the leading DT developer for the manufacturing company and expanding the value network with partner organizations to add value to the DT system for customers.

The various stakeholders can add value to a DT ecosystem in several ways. For example, [33] identified that intelligent robots could streamline manufacturing processes to operate faster and more accurately in future smart manufacturing environments. Dhanabalan and Sathish [34] studied how AI-based robots would improve manufacturing processes by providing enhanced monitoring and auto-correction of processes when robots are present in the manufacturing environment. These intelligent robots are capable of self-configuring, self-adjusting, and self-optimizing through the data from DTs, making manufacturing processes more resilient. However, reaping the benefits inevitably requires working in DT ecosystems. In line with prior literature [16], interconnected stakeholders play an important role in providing better products and services to customers and in deriving value in ecosystems.

Ekman et al. [35] suggested that the co-creation of value within the ecosystem depends on the actors within it, and further stated that over time, stakeholders may discover new ways of utilizing services within the ecosystem so that new forms of value are created. Identifying the possible stakeholders in the design stage will help industries develop a better ecosystem while satisfying the needs of these stakeholders.

Different actors in a DT ecosystem have different objectives, and their decision-making criteria can differ from one another, as Tsujimoto et al. [16] found in their study. Furthermore, the decisions and behaviors of one stakeholder affect others, ultimately influencing the ecosystem's evolution or decline. As evident in the results of the present study, the different stakeholders have different goals when they engage collaboratively in the development of DTs. Therefore, working in DT ecosystems may ultimately be the decision of key stakeholders (e.g., customer or platform owner).

Table 6
Mapping the potential challenges faced by stakeholders (see Fig. 2).

Challenge	Stakeholder/Stakeholder Category
Interoperability	Primary stakeholders, secondary stakeholders
Cybersecurity	All stakeholders
Data Governance and Management	Manufacturers, suppliers, subcontractors, secondary stakeholders, services, tertiary stakeholders
Data Sharing/Data Fusion	Manufacturers, suppliers, subcontractors, secondary stakeholders, services, tertiary stakeholders
Orchestration	Manufacturers, suppliers, subcontractors, secondary stakeholders, services, tertiary stakeholders
Gaining Stakeholder Commitment	Manufacturers, tech companies
DT expertise	Tech companies, manufacturers, physical asset providers

5.2. Challenges in DT ecosystems (answer to RQ2)

Upon the identification of the possible obstacles, as outlined in the SLR and the empirical evidence collected through the interviews, it is evidently revealed that additional potential challenges can be faced in an ecosystem-focused DT development. While the existing literature on DT ecosystems has focused on data-related issues, including data governance, data sharing, and data fusion [15,25,29], our empirical results highlight the importance of addressing the entire process of data management in DT ecosystems, as the spectrum of challenges reaches beyond mere data-related issues.

Apart from the challenges mentioned in the literature, such as interoperability, cybersecurity, and orchestration [15,24,28], this study identifies new technical challenges, such as gaining stakeholder commitment and DT expertise, which were not identified in previous literature. In addition to these, our empirical evidence has identified non-technical challenges that were also not evident in the SLR. These involve justifying usability and need and the potential value generated from the DT ecosystem for stakeholders (i.e., gaining stakeholder commitment) and sourcing skilled labor for complex developments (i.e., DT expertise). Thus, this study has identified new technical and non-technical challenges in a DT ecosystem. These new challenges are increasingly critical, especially at the ecosystem level. The mapping of these challenges with the stakeholders identified in our earlier study [7] is presented in Table 6.

5.3. Guidelines to address the challenges

The existing literature on DT ecosystems mostly identifies solutions to security and data governance challenges. These solutions presented in the literature use blockchain technology to eliminate security challenges and ensure encryption [24,29]. Thus, the literature does not discuss solutions broadly for the possible challenges identified in the SLR. The present research suggests standardization through the use of frameworks such as Open Platform Communications Unified Architecture (OPC UA) and ISO 23247 [36] as one of the solutions for the challenges identified in the SLR and the interviews, such as data governance, interoperability, data sharing/fusion, and data management. The use of OPC UA and ISO 23247 was not presented in any of the studies considered for the SLR. ISO-23247 offers a standard digital twin development framework that helps partners in the digital twin ecosystem to analyze requirements and use a common language when communicating with each other. The framework also facilitates the management and sharing of data from different parts of digital twin systems among ecosystem partners. Solheim and Powell [37] suggested that related stakeholders can use OPC UA specification to achieve standardization. An interviewee also suggested this as a possible way of implementing standardization.

Table 7
Guidelines for dealing with challenges.

Challenge	Guidelines for Dealing with Challenges
Interoperability	OPC UA and ISO 23247 can provide frameworks and technologies for the effective interoperability of digital twins.
Cybersecurity	Data encryption during transmission and use of blockchain-based solutions
Data Management, Data sharing/Fusion	Useful data management and sharing strategies for digital twins can be provided by OPC UA and ISO 23247.

Furthermore, the above-identified stakeholder in the study, platform integration service providers, could perform a major role in standardization within the ecosystem. They could develop APIs, cloud-based automated tools, and environments that could be used by ecosystem partners, such as tech companies and physical asset suppliers, in developing their applications and systems to help these systems integrate easily and reduce the challenges of interoperability, data governance, and so on. Based on the study results, the following guidelines, as shown in Table 7, can be provided in dealing with different challenges within the ecosystem. Gaia-X and Dataspaces are initiatives that aim to establish secure data infrastructures, which can effectively tackle the issues of data management, sharing, and privacy across various industries and domains [38]. These initiatives can improve collaboration between ecosystem partners for data management and sharing while adhering to regulatory requirements.

5.4. DT ecosystem definition

During the SLR process, we identified several DT ecosystem studies. However, we could not find a reasonable definition of DTE. Based on the knowledge gained from our study, we defined it as follows:

> *A digital twin ecosystem is a data-driven network of interconnected instances of a digital twin or different digital twins, along with different organizational and individual stakeholders, that will create value for one another, enabled by new technologies.*

This definition represents four main characteristics of the DT ecosystem: data driven, interconnected DTs, stakeholders who create value, and technologies enabled by technologies. By its very nature, this definition of a DT ecosystem has broadened earlier definitions by including organizational and individual stakeholders. These individual stakeholders can also include non-living individuals, such as AI robots. As such, this definition represents the above-identified stakeholders in this study. It also emphasizes that the network of interconnected DTs is data driven, which is a fundamental characteristic of DTs.

This study is significant because it provides a better picture of stakeholders and their requirements in a DT ecosystem by identifying new stakeholders, and thus provides better insights into the DT ecosystem. This definition would also be useful in future research and in the domain of DT ecosystems, as it provides a holistic view of them that would shed light on their broad nature.

5.5. Implications on research

The overall significance of this study for research and academia is that it comprehensively examines collaboration related to DT ecosystems and identifies potential challenges related to DT development, particularly when moving from a value chain-focused DT to an ecosystem-focused DT. Most of the literature on DT ecosystems has delved into the application of different technologies to build DT ecosystems or possible scenarios for building DT ecosystems [39]. These studies do not comprehensively address the aforementioned aspects considered in the present research. The current work also provides a useful understanding of how stakeholders can address different challenges. This means that the findings can be used by future research to understand the relevance of different challenges to different stakeholders or how different stakeholders could participate in eliminating such challenges. Notably, the results represent the perspectives of both academia and industry, thereby providing a comprehensive view of the studied phenomenon and new knowledge to the literature. Additionally, the new definition of DT ecosystem presented in this study would be useful in future research to identify the broad horizons of the DT ecosystem domain.

5.6. Implications on practice

For a DT to operate efficiently, identifying the potential challenges that may arise when operating on a broader scale, such as on an ecosystem level, is critical. This study allows companies that engage in collaboration and develop a DT ecosystem to gain an initial understanding of these possible challenges, which, in turn, enables them to derive solutions and thus realize the intended benefits. Our study also highlights important non-technical challenges that need to be considered by ecosystem partners before moving to the ecosystem level in DT collaboration and development. This study provides valuable insights for the identified stakeholders in developing more coherent DT ecosystems by adopting the necessary solutions to overcome difficulties.

Furthermore, this study outlines the importance of standardization, one of the best possible solutions for challenges in a DT ecosystem. For standardization, industries could use OPC UA specification, for example, to help address the wide range of challenges relating to interoperability, data governance, data sharing, and data management. As such, this research helps tech companies identify challenges and provide possible solutions. Practitioners could use these guidelines to deal with difficulties in the DT ecosystem, or they could use them in the design stage to develop possible solutions.

Overall, this study addresses a wide range of perspectives in the domain of the DT ecosystem, which can guide practitioners in the design, development, and maintenance of the DT ecosystem. Practitioners would vastly benefit from this study, as it would help them manage a healthy ecosystem better while deriving value for stakeholders in the ecosystem.

5.7. Future research

This study provides several research avenues for the future. First, it identifies different possible challenges that could be encountered when designing and developing DTs in ecosystems. In future research, these identified challenges can be considered when deploying a DT ecosystem. This approach may help pinpoint potential technologies for overcoming such challenges. Second, future research could also study the identified challenges in greater depth to implement appropriate solutions, thus significantly impacting the development of this domain of the DT ecosystem.

5.8. Emerging concepts in DT ecosystems

5.8.1. Interdisciplinary stakeholder collaboration

Collaboration among different stakeholders from different fields can play a crucial role in enhancing the understanding of digital twin development and evolution. This collaboration among experts from various disciplines also helps ecosystems remain effective and sustainable as technology evolves and new challenges emerge. When people from different backgrounds collaborate, they can better predict and deal with these changes.

5.8.2. Dynamic digital twin: From adaptability to evolvability

Our work underpins the dynamism of DTs for human–machine collaborative work. DTs must be updated with all of the physical systems' adaptive changes and modifications. As a result, the simulation generated during operations matures in terms of both data and physical device connectivity. This can be seen as the dynamic adaptability of DT systems for local optimization. The current research is mainly focused on the adaptability of closed DT systems at the local level. We propose to extend the focus also to incorporate the DT system's evolvability. Through the case of Digital Twin B, we determined that data- and knowledge-centric DTs need to be implementable at the ecosystem level in order to be economically viable.

5.8.3. Low-code/no-code development

The implementation of low-code/no-code is an innovative approach for building applications on top of DT systems. This approach facilitates the evolvability of a DT ecosystem, making the DT systems adaptable and integrable by ecosystem partners. This approach is also cost saving and shortens the DT application development cycle. Moreover, DT application developers can further refine the low-code or no-code approach to use open-source software packages, frameworks, and APIs in order to speed up the application development process, further supporting the shift from adaptability to the evolvability of Digital Twin B systems.

5.8.4. Blueprint for data-centric service development

Data-centric service development refers to the approach in which data emerging from various sources, such as sensors, are centered and integrated into a unified model within data storage. In this setup, systems and services are developed with a primary focus on these centralized data, which leads to data-centric systems and service development (e.g., Digital Twin A). Thus, the stakeholders in an ecosystem need to decide on the practices that enable many parties (e.g., external software development) to be integrated for application and service development on the extracted data and knowledge sources. Examples of such practices/solutions include information security and confidentiality assurance and the interfaces provided to developers. As complex systems, DT systems are tailored to specific customers or usage scenarios requiring different configurations and/or components. The potential research topics that emerge are as follows:

- Data models and standards: The current emphasis is on broadening the reach of existing data models and deploying them. Different specifications exist for each component in each layer of the software and hardware systems. The interoperability between these technical layers is a key next step. The end results should be stable, reliable, and effective for industry-wide data sharing. Collaboration with other domain standards to promote interoperability is a potential goal for DT ecosystems.
- Data management and integration: It is particularly important to concentrate on semantic precision when mapping information between the areas of data science and information management (i.e., the data must be combined on a common basis to maintain their true meaning). Standards that enable data integration, with a focus on semantic precision, make it easier to build transparent and long-lasting data systems that would significantly benefit stakeholders.
- Data protection and privacy: The emergence of DT ecosystems has led to concerns about data hosting, ownership, and stakeholder privacy, as they bridge information gaps and allow data to flow seamlessly between operational lifecycles. Solutions can include identifying data hosting, protection, and privacy criteria, creating (joint) agreements for ecosystem governance at the industrial level, and harnessing a community-based approach.

In summary, data-centric service development is an approach in which services and applications are designed around centralized data, ensuring seamless integration, interoperability, and effective management, all while maintaining security and privacy standards.

5.9. Study validity

This section discusses validity in terms of construct validity, external validity, and reliability using the guidelines by Yin [40].

- Construct validity means taking the right measures to examine the phenomenon under study and taking full precaution during data collection so that the collected data align with the given research questions. A SLR collects data from existing academic research in the field. We designed our interview guide to focus on our research objectives. During the interview selection, experienced people within the organization were chosen for the interviews, keeping the following aspects in mind: their knowledge of the DT ecosystem and their DT development experience in ecosystem environments. We also reviewed the interview guide along with a company representative to ensure that the participants would properly understand the questions. The results of the SLR and the interviews were comprehended to answer the research questions. Therefore, construct validity can be seen as being successfully achieved in the study's theoretical part and the empirical study.
- External validity refers to how findings can be generalized outside the investigated cases. The results of this study are limited to two DT system units. However, for companies using similar large-scale frameworks in their DT developing organizations, the results of this study can be applicable and useful in their contexts. The literature review and the empirical study present insights into stakeholder collaboration and the challenges that will be encountered when moving to the DT ecosystem level from different perspectives. As the study is empirical, this reduces the possible opportunities for bias in the results that would have prevailed in the findings of the SLR. As such, the results can be applied to similar contexts.
- Reliability is concerned with how data and the analysis depend on the specific researchers involved. During data collection, the first and second authors designed the interview guide to obtain information on the DT ecosystem from various viewpoints. The fourth author audited and supervised the entire process. The summary of the findings was sent to the interviewees via email, who then validated the content. A company representative further reviewed the findings to validate the data.
  Furthermore, this study described the research process, both the SLR and the empirical study, and how the data were analyzed to answer the research questions. The questionnaire used in the interview was also presented in the methodology section. However, there is a possibility that another researcher might identify different results, as the data gathered from semi-structured interviews can change depending on the interactions between the interviewer and the interviewees, the situation, and the accumulated knowledge of the interviewees at the time. As such, there is a possibility that the results of the empirical study could change if this research is repeated.

5.10. Study limitations

The development of a DT ecosystem is still in its infancy. When collecting data from the SLR, most DT ecosystems were still at the conceptual level and did not represent the actual implemented DT ecosystems. This could have affected the results of the SLR. We also acknowledge the limitation of conducting only eight expert interviews. However, our semi-structured interviews with open-ended questions, allowed us to explore deeper into the topic. In the future, it would be beneficial to include a wider variety of expert interviews to gain a more comprehensive understanding of the subject matter. Also, as the ecosystem DT perspective is not well defined and may convey a different meaning to different people, the participants might have misinterpreted the questions, and their responses could have also been based on conjecture because of their limited experiences with DT ecosystems. Therefore, their answers could be subjective and biased, depending on their knowledge and expertise. Finally, the interviews were conducted in English. As such, some participants experienced difficulties articulating their ideas in English. This may have affected the quality of the data gathered from the interviews.

6. Conclusion

This comprehensive study delved into the intricate dynamics of DT ecosystems, emphasizing the pivotal role of stakeholder collaboration and the challenges and solutions inherent in the transition from a value chain-focused DT to an ecosystem-focused DT. The research underscored the multifaceted roles and unique requirements of stakeholders, from manufacturers to intelligent robots, in the DT ecosystem. Through a meticulous examination of the academic literature and empirical evidence, the study identified technical and non-technical challenges that industries might face when scaling up their DT operations. Notably, the research highlighted the importance of standardization as a potential solution to many of these challenges.

The study also introduced an expanded definition of the DT ecosystem, emphasizing its data-driven nature, interconnected DTs, value-creating stakeholders, and technology-enabling capabilities. This definition broadens the understanding of DT ecosystems and provides a holistic view that encompasses both organizational and individual stakeholders, including AI robots.

From a practical standpoint, the insights derived from this study offer invaluable guidance for industries aiming to design, develop, and maintain DT ecosystems. The research provides a roadmap for navigating challenges and harnessing the potential benefits of DT ecosystems. Moreover, the study's findings serve as a foundation for future research endeavors, offering numerous avenues for further exploration, from the evolvability of DT systems to data-centric service development.

Finally, this study contributes significantly to DT ecosystems' academic and industrial domains, offering a comprehensive perspective that bridges the gap between theory and practice. Through its findings, the research paves the way for a more collaborative, efficient, and value-driven future for DT ecosystems.

CRediT authorship contribution statement

Nirnaya Tripathi: Writing – review & editing, Writing – original draft, Visualization, Project administration, Methodology, Investigation, Data curation, Conceptualization. **Heidi Hietala:** Writing – review & editing, Writing – original draft. **Yueqiang Xu:** Writing – original draft, Investigation. **Reshani Liyanage:** Methodology, Investigation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Nirnaya Tripathi reports financial support was provided by University of Oulu. Nirnaya Tripathi reports a relationship with University of Oulu that includes: employment.

Data availability

The data that has been used is confidential.

Acknowledgments

This research was supported by the Oxilate project, Business Finland. We thank Tero Päivärinta, Finland for the support.

Appendix A. Primary articles description

ID, Reference	Description
[P1] [24]	Propose a DTwins ecosystem that is built upon Digital Twins (DT) characteristics, personal health system requirements, ubiquitous biofeedback, and a cyber–physical reference model. Its primary objective is establishing the groundwork for a practical DT system in preventive healthcare.
[P2] [25]	The purpose of the article is to introduce a framework called “Smart Rural Area Data Infrastructure” (SRADI) to interconnect resources among stakeholders and platforms in an ecosystem while enabling the integration of both historical and real-time information for physical assets in agriculture.
[P3] [26]	This paper presents a blockchain-based model for collaborative digital twin ecosystems, ensuring security and privacy. It follows “X-by-design” and “X-as-a-service” principles, exemplified through case studies of storing digital twin operations and transactions on the blockchain, promising improved security, accountability, and integrity in decentralized systems for the “Any 4.0” era.
[P4] [27]	This paper presents a service-oriented hybrid digital twin and digital thread platform for enhancing collaboration in a smart product and service ecosystem (SPSS). The approach incorporates crowd-/service-sourcing and Internet of Things (IoB) to support advanced manufacturing services. It introduces a conceptual model, system realization techniques, and a case study to demonstrate its feasibility in the heating industry.
[P5] [28]	This paper discusses the concept of digital twins (DTs) in urban water systems (UWS), categorizing them into living and prototyping DTs. It emphasizes the importance of multifunctional, updateable DTs connected to real-time observations and simulation models. The study clarifies DT terminology for UWS and outlines steps for creating DTs using digital ecosystems and open data standards.
[P6] [29]	The paper discusses the growing importance of Digital Twins in industrial digitization and the need for secure data sharing. It proposes an owner-centric decentralized sharing model using Decentralized Applications, addressing data integrity and confidentiality. The prototype, EtherTwin, overcomes implementation challenges and is evaluated for industry use cases, with a focus on investigating the suitability of blockchain technology in the Digital Twin sharing ecosystem.
[P7] [19]	This paper explores the various perspectives and conceptions of Digital Twins, emphasizing its significance as a technology trend. It outlines a vision for a next-generation Digital Twin in mechatronics systems, focusing on semantic technologies, model transfer, and service generation within an ecosystem. The vision is illustrated through a lemonade production example, demonstrating the potential for Digital Twins to enable innovative services throughout the product lifecycle.
[P8] [6]	The paper highlights the growing importance of Digital Twins (DT) in manufacturing. It discusses various technological visions for DT and emphasizes the need for a sociotechnical design approach to create DT software specific to users and their evolving needs. The paper introduces a vision for a DT-based Digital Platform supporting product design and lifecycle management within an ecosystem of twin-driven product-service systems.
[P9] [30]	This paper addresses the challenges of information management (IM) in the aircraft manufacturing industry using Digital Twin (DT) technology. It introduces an IM framework tailored to the sector, highlighting key phases and elements. The framework enhances information management and paves the way for further research opportunities in the DT domain within the aircraft lifecycle.
[P10] [15]	This paper explores the software development perspective on digital twins in the context of Industry 4.0, emphasizing the need for an ecosystem view. It presents a framework that considers the scope of the digital twin software platform, life-cycle phases, and the level of integration with the physical system. The paper suggests further research to build a comprehensive research agenda based on a systematic literature review, addressing key research questions within this framework.

Appendix B. Interview questions

Questions
1. What is your role in the organization, and how is your work related to digital twins?
2. Could you please explain the products and services that you offer to your customers?
3. What is a digital twin in your opinion?
4. Do you develop or utilize digital twins?
a. Who are your key stakeholders?
b. How do you manage your connections with key stakeholders?
5. Do you cooperate with other companies, such as suppliers and customers, in developing your digital twin?
6. Do you believe in ecosystem-focused digital twins, and what do you think of them?
7. Could you please explain your organization’s role within this ecosystem?
8. What challenges do you think will be encountered in an ecosystem focused on digital twin development?
9. What potential solutions can be adopted when dealing with challenges in digital twin ecosystems?

References

[1] Peter Augustine, The industry use cases for the digital twin idea, in: Advances in Computers, vol. 117, (no. 1) Elsevier, 2020, pp. 79–105.
[2] Anna Hyre, Gregory Harris, John Osho, Minas Pantelidakis, Konstantinos Mykoniatis, Jia Liu, Digital twins: Representation, replication, reality, and relational (4rs), Manuf. Lett. 31 (2022) 20–23.
[3] Teng Wang, Zheng Liu, Min Liao, Nezih Mrad, Life prediction for aircraft structure based on Bayesian inference: Towards a digital twin ecosystem, in: Annual Conference of the PHM Society, vol. 12, (no. 1) 2020, p. 8.
[4] Shohin Aheleroff, Xun Xu, Ray Y. Zhong, Yuqian Lu, Digital twin as a service (dtaas) in industry 4.0: An architecture reference model, Adv. Eng. Inform. 47 (2021) 101225.
[5] Concetta Semeraro, Mario Lezoche, Hervé Panetto, Michele Dassisti, Digital twin paradigm: A systematic literature review, Comput. Ind. 130 (2021) 103469.

[6] Henrique Diogo Silva, Miguel Azevedo, António Lucas Soares, A vision for a platform-based digital-twin ecosystem, IFAC-PapersOnLine 54 (1) (2021) 761–766.

[7] Reshani Liyanage, Nirnaya Tripathi, Tero Päivärinta, Yueqiang Xu, Digital twin ecosystems: Potential stakeholders and their requirements, in: Software Business: 13th International Conference, ICSOB 2022, Bolzano, Italy, November 8–11, 2022, Proceedings, Springer, 2022, pp. 19–34.

[8] Barbara Kitchenham, Stuart Charters, et al., Guidelines for performing systematic literature reviews in software engineering, 2007.

[9] Eric Armengaud, Christoph Sams, Georg Von Falck, Georg List, Christian Kreiner, Andreas Riel, Industry 4.0 as digitalization over the entire product lifecycle: Opportunities in the automotive domain, in: Systems, Software and Services Process Improvement: 24th European Conference, EuroSPI 2017, Ostrava, Czech Republic, September 6–8, 2017, Proceedings 24, Springer, 2017, pp. 334–351.

[10] John Lee, Ian Cameron, Maureen Hassall, Improving process safety: What roles for digitalization and industry 4.0? Process Saf. Environ. Prot. 132 (2019) 325–339.

[11] Michael Grieves, John Vickers, Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems, Transdiscipl. Perspect. Complex Syst.: New Findings Approaches (2017) 85–113.

[12] R. Rosen, G. Von Wichert, G. Lo, et al., About the importance of autonomy and digital twins for the future of manufacturing, IFAC-PapersOnLine (2015).

[13] Luiz Fernando C.S. Durão, Sebastian Haag, Reiner Anderl, Klaus Schützer, Eduardo Zancul, Digital twin requirements in the context of industry 4.0, in: Product Lifecycle Management To Support Industry 4.0: 15th IFIP WG 5.1 International Conference, PLM 2018, Turin, Italy, July 2-4, 2018, Proceedings 15, Springer, 2018, pp. 204–214.

[14] W. Kritzinger, M. Karner, G. Traar, J. Henjes, W. Sihn, Digital twin in manufacturing: A categorical literature review and classification, IFAC-PapersOnLine 51 (2018) 1016–1022.

[15] Yueqiang Xu, Tero Päivärinta, Pasi Kuvaja, Digital twins as software and service development ecosystems in industry 4.0: Towards a research agenda, in: Big Data and Security: First International Conference, ICBDS 2019, Nanjing, China, December 20–22, 2019, Revised Selected Papers 1, Springer, 2020, pp. 51–64.

[16] Masaharu Tsujimoto, Yuya Kajikawa, Junichi Tomita, Yoichi Matsumoto, A review of the ecosystem concept—Towards coherent ecosystem design, Technol. Forecast. Soc. Change 136 (2018) 49–58.

[17] Marco Iansiti, Roy Levien, Strategy as ecology, Harvard Bus. Rev. 82 (3) (2004) 68–78.

[18] Xiufeng Yan, Construction of digital twin ecosystem for coal-fired generating units, J. Phys.: Conf. Ser. 1748 (5) (2021) 052037.

[19] Roland Rosen, Jan Fischer, Stefan Boschert, Next generation digital twin: An ecosystem for mechatronic systems? IFAC-PapersOnLine 52 (15) (2019) 265–270.

[20] Minas Pantelidakis, Konstantinos Mykoniatis, Jia Liu, Gregory Harris, A digital twin ecosystem for additive manufacturing using a real-time development platform, Int. J. Adv. Manuf. Technol. 120 (9–10) (2022) 6547–6563.

[21] Roberto Minerva, Gyu Myoung Lee, Noel Crespi, Digital twin in the IoT context: A survey on technical features, scenarios, and architectural models, Proc. IEEE 108 (10) (2020) 1785–1824.

[22] Martin Ivarsson, Tony Gorschek, A method for evaluating rigor and industrial relevance of technology evaluations, Empir. Softw. Eng. 16 (2011) 365–395.

[23] You Zhou, He Zhang, Xin Huang, Song Yang, Muhammad Ali Babar, Hao Tang, Quality assessment of systematic reviews in software engineering: A tertiary study, in: Proceedings of the 19th International Conference on Evaluation and Assessment in Software Engineering, 2015, pp. 1–14.

[24] Abdulmotaleb El Saddik, Hawazin Badawi, Roberto Alejandro Martinez Velazquez, Fedwa Laamarti, Rogelio Gámez Diaz, Namrata Bagaria, Juan Sebastian Arteaga-Falconi, Dtwins: A digital twins ecosystem for health and well-being, IEEE COMSOC MMTC Commun. Front. 14 (2) (2019) 39–43.

[25] David Gackstetter, Mandana Moshrefzadeh, Thomas Machl, Thomas H Kolbe, Smart Rural Areas data infrastructure (SRADI)–An information logistics framework for digital agriculture based on open standards, 2021, 41. GIL-Jahrestagung, Informations-und Kommunikationstechnologie in kritischen Zeiten.

[26] Alper Kanak, Niyazi Ugur, Salih Ergun, A visionary model on blockchain-based accountability for secure and collaborative digital twin environments, in: 2019 IEEE International Conference on Systems, Man and Cybernetics, SMC, IEEE, 2019, pp. 3512–3517.

[27] Xiaojing Niu, Shengfeng Qin, Integrating crowd-/service-sourcing into digital twin for advanced manufacturing service innovation, Adv. Eng. Inform. 50 (2021) 101422.

[28] Agnethe N. Pedersen, Morten Borup, Annette Brink-Kjær, Lasse E. Christiansen, Peter S. Mikkelsen, Living and prototyping digital twins for urban water systems: Towards multi-purpose value creation using models and sensors, Water 13 (5) (2021) 592.

[29] Benedikt Putz, Marietheres Dietz, Philip Empl, Günther Pernul, Ethertwin: Blockchain-based secure digital twin information management, Inf. Process. Manage. 58 (1) (2021) 102425.

[30] Sumit Singh, Essam Shehab, Nigel Higgins, Kevin Fowler, John A. Erkoyuncu, Peter Gadd, Towards information management framework for digital twin in aircraft manufacturing, Procedia CIRP 96 (2021) 163–168.

[31] Daniela S. Cruzes, Tore Dyba, Recommended steps for thematic synthesis in software engineering, in: 2011 International Symposium on Empirical Software Engineering and Measurement, IEEE, 2011, pp. 275–284.

[32] Yuqian Lu, Chao Liu, I Kevin, Kai Wang, Huiyue Huang, Xun Xu, Digital twin-driven smart manufacturing: Connotation, reference model, applications and research issues, Robot. Comput.-Integr. Manuf. 61 (2020) 101837.

[33] Mohd Javaid, Abid Haleem, Ravi Pratap Singh, Rajiv Suman, Substantial capabilities of robotics in enhancing industry 4.0 implementation, Cognit. Robot. 1 (2021) 58–75.

[34] T. Dhanabalan, A. Sathish, Transforming Indian industries through artificial intelligence and robotics in industry 4.0, Int. J. Mech. Eng. Technol. 9 (10) (2018) 835–845.

[35] Peter Ekman, Randle D. Raggio, Steven M. Thompson, Service network value co-creation: Defining the roles of the generic actor, Ind. Mark. Manag. 56 (2016) 51–62.

[36] JaeSeung Song, Franck Le Gall, Digital twin standards, open source, and best practices, in: The Digital Twin, Springer, 2023, pp. 497–530.

[37] Anja Bottinga Solheim, Daryl John Powell, A learning roadmap for digital lean manufacturing, in: Advances in Production Management Systems. Towards Smart and Digital Manufacturing: IFIP WG 5.7 International Conference, APMS 2020, Novi Sad, Serbia, August 30–September 3, 2020, Proceedings, Part II, Springer, 2020, pp. 417–424.

[38] Boris Otto, The evolution of data spaces, in: Designing Data Spaces: The Ecosystem Approach to Competitive Advantage, Springer International Publishing Cham, 2022, pp. 3–15.

[39] Cordelia Mattuvarkuzhali Ezhilarasu, Zakwan Skaf, Ian K. Jennions, Understanding the role of a digital twin in integrated vehicle health management (IVHM), in: 2019 IEEE International Conference on Systems, Man and Cybernetics, SMC, IEEE, 2019, pp. 1484–1491.

[40] Robert K. Yin, Case Study Research: Design and Methods, vol. 5, sage, 2009.