Data & Knowledge Engineering 151 (2024) 102304

Contents lists available at ScienceDirect

Data & Knowledge Engineering

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

Artificial intelligence in digital twins—A systematic literature review

Tim Kreuzer *, Panagiotis Papapetrou, Jelena Zdravkovic

Department of Computer and Systems Sciences, Stockholm University, Borgarfjordsgatan 12, Kista, 16455, Sweden

A R T I C L E I N F O

Keywords:
Artificial intelligence
Digital twin
Machine learning
Literature review
Business intelligence
Data mining

A B S T R A C T

Artificial intelligence and digital twins have become more popular in recent years and have seen usage across different application domains for various scenarios. This study reviews the literature at the intersection of the two fields, where digital twins integrate an artificial intelligence component. We follow a systematic literature review approach, analyzing a total of 149 related studies. In the assessed literature, a variety of problems are approached with an artificial intelligence-integrated digital twin, demonstrating its applicability across different fields. Our findings indicate that there is a lack of in-depth modeling approaches regarding the digital twin, while many articles focus on the implementation and testing of the artificial intelligence component. The majority of publications do not demonstrate a virtual-to-physical connection between the digital twin and the real-world system. Further, only a small portion of studies base their digital twin on real-time data from a physical system, implementing a physical-to-virtual connection.

1. Introduction

Digital twins are virtual replicas of real-world systems that accurately reflect the system's behavior, with the goal of achieving functionalities like automatic fault diagnosis or run-time monitoring. A Digital Twin (DT) can represent an object, such as a smart building, or a process, such as a production process. Following the definition of Jones et al. [1], a DT is characterized by a physical-to-virtual connection, the data stream that replicates the real-world system. Additionally, the inverse, a virtual-to-physical connection, closes the loop between virtual and physical space, allowing the physical system to benefit from the output of the digital twin. A digital twin should have the ability to process various kinds of data, including real-time data, for providing real-time monitoring of the physical system it represents - this is essential for immediate detection of the original system's behavior, its critical events or anomalies, and response to them.

Artificial Intelligence (AI) and Machine Learning (ML) solutions have become an important part of research in many disciplines in recent years. AI has been shown to outperform humans in several different tasks, such as reconstructing brain circuits [2], playing strategy games [3], and predicting protein structure [4]. Due to the wide applicability of AI solutions, different systems can benefit from the integration of an AI component for predictive functionality.

Since both AI and DT systems require data to function, it is a logical step to integrate them. Integrations of AI and DT, creating an AI-DT system, have been proposed in different application domains, where the AI component makes predictions based on data stemming from the DT. As some examples, Fahim et al. [5] describe a digital twin of a wind turbine using machine learning to forecast the system's energy production. Xiong et al. [6] propose a solution for predictive maintenance of an aircraft engine,

* Corresponding author.
E-mail addresses: tim.kreuzer@dsv.su.se (T. Kreuzer), panagiotis@dsv.su.se (P. Papapetrou), jelenaz@dsv.su.se (J. Zdravkovic).

https://doi.org/10.1016/j.datak.2024.102304
Received 18 December 2023; Received in revised form 25 March 2024; Accepted 29 March 2024
Available online 3 April 2024

combining a DT of the engine with an LSTM model [7] serving as the ML component. In all such contributions, the Digital Twin derives significant advantages from the application of ML techniques, and through the utilization of vast historical data and pertinent AI algorithms, attains the capability to enhance the accuracy of its predictions. This is achieved by making use of the analytical power of ML, which in turn leverages information and patterns embedded in the historical data to refine predictions, leading to more precise and reliable DT outputs.

This study systematically searches and investigates the literature on the intersection of digital twins and artificial intelligence. After the search, relevant studies are referred to give an overview of the state of the art. The results provide researchers with a summary of the related work, serving as foundational knowledge for future work in the field. Based on this, we identify a number of relevant insights and gaps within current research, suggesting directions for future work.

In recent research, some literature reviews on artificial intelligence and digital twins have been conducted [8–10]. However, although similar in the overarching topic, the focus, search scope of these reviews, and thereby identified research gaps, differ substantially from our review.

Kritzinger et al. [11] define three levels of integration for digital twins: *Digital models*, which have a manual information flow between the physical and the virtual system in both directions; *Digital shadows*, which automate the physical-to-virtual data flow, but still have a manual virtual-to-physical feedback loop; *Digital twins* are the most advanced level of system integration, providing a bidirectional automated data flow between real-world system and digital twin. In this paper, we follow these definitions for digital model, digital shadow, and digital twin.

This paper is structured as follows: Section 2 provides the background for this work and characterizes related literature reviews. Section 3 describes the methodology employed for our literature review, describing the search process and the research questions investigated. Section 4 presents the results of the literature search, characterizing the state of the art, and general findings regarding past research in the field of AI in digital twins. Further, results specific to the research questions are presented. Section 5 identifies research gaps and discusses the findings. Lastly, Section 6 concludes the paper, summarizing the contributions.

2. Background

Jones et al. [1] have conducted a systematic literature review on digital twins, identifying characteristics of a digital twin and highlighting gaps in the research field. The authors do not explicitly mention machine learning or artificial intelligence, however, concepts such as predictive maintenance and advanced control systems are emphasized as components of a digital twin. Semeraro et al. [12] have performed a review on the paradigm of digital twins, with a focus on the definition of a digital twin and its application domains, based on text mining techniques. While their research focuses on the components of a digital twin, artificial intelligence is not acknowledged in the article.

Schmid and Winkler [8] have performed a literature review on the combination of AI methods and digital twins of production systems. They review common challenges encountered in related work and propose a framework combining human interaction and automated components utilizing AI in a production system. A systematic literature review on the role of artificial intelligence, machine learning, and big data within digital twins has been carried out by Rathore et al. [9]. The authors group the existing literature by application domain, characterizing the use-case of the analyzed studies, as well as the machine learning solution employed. Further, they focus on tools supporting the creation of digital twins, providing a reference architecture of their design. While Rathore et al. [9] focus on application domains and the tools supporting digital twin development, this paper is focused on the tasks of artificial intelligence within digital twins as well as the modeling approaches pursued in existing work.

Lim et al. [13] survey state-of-the-art digital twin techniques, sampling exclusively journal articles. Their review focuses on the lifecycle stages of a DT system, positioning past research work by lifecycle stage. Further, the authors identify the integration of big data and digital twins as a future perspective for DT research, stating that the combination can improve decision-making support and improve the quality of simulations within the DT. Bartsch et al. [10] have conducted a literature review on the application of artificial intelligence methods in digital twins for additive manufacturing. They survey a small set of papers specific to digital twins in additive manufacturing as well as a separate set of papers focusing on artificial intelligence in additive manufacturing. The authors state that there is a need for the integration of AI methods with digital twins, however, their analysis does not connect the two research areas.

The previously given definition of a digital twin by Kritzinger et al. [11] is aligned with the definition by Grieves [14], who first mentioned digital twins. Grieves states that a digital twin requires a physical system, a virtual counterpart, and a two-way connection between them. Additionally, his definition highlights the real-time application scenario as paramount, as digital twins provide the most value when applied in a real environment, making an impact on a live system, rather than being applied in a lab setting on synthetic data. Fig. 1 shows a schematic model of a digital twin, adapted from Grieves [15], with the addition of an AI component, showing the bidirectional information flow between the DT and the real-world system. The real-world system, which is the physical representation, contains sensors such as a temperature sensor and actors, which are active components that the DT can control, such as a ventilation system. The AI component is a part of the digital twin, which is a virtual representation of the physical system. The AI component fulfills predictive tasks, such as performing a forecast based on real-time temperature data. The digital twin has additional capabilities, such as data analysis, decision-making, and scenario simulation, which depend on the use case and are summarized in the figure as *DT Capabilities*. The physical and virtual representations are bidirectionally connected, where the virtual-to-physical connection represents a feedback loop from the digital twin, which sends feedback to the physical system based on its internal processing and decision-making, while the physical-to-virtual connection is the data stream supplying the DT based on measurements from the real-world system.

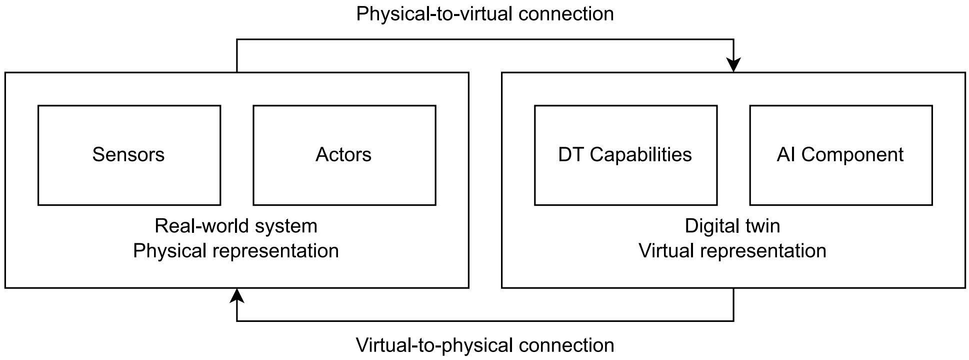

Fig. 1. Schematic model of a digital twin with an AI component.

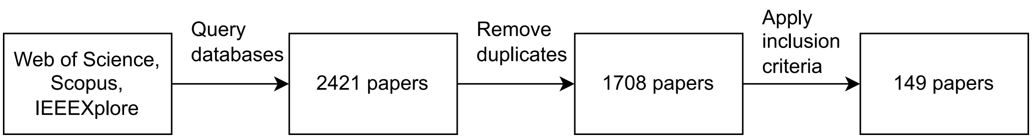

Fig. 2. Literature search process.

In this paper, AI is used as an overarching term describing the research field, which encapsulates machine learning and deep learning (DL). Deep learning is a subfield of artificial intelligence, where, following the definition of LeCun et al. [16], computational models learn representations of data with multiple levels of abstraction. In practice, deep learning uses different types of neural networks in combination with non-linear activation functions to pursue different tasks, such as regression or classification. Reinforcement learning (RL) is a process where an agent learns to solve a problem by trial-and-error [17], which is typically encountered in optimization scenarios. AI algorithms are often black-box models which are not inherently understandable for the user. The field of explainable AI is concerned with the interpretation of AI models for human users.

Feedforward neural networks are the foundational model architecture of deep learning. They have been used in practice since the inception of the backpropagation algorithm [18], allowing them to learn complex, nonlinear problems. Convolutional neural networks (CNNs) [19] are a special type of neural networks, that subsample the input data with spatial or temporal convolutions, allowing them to capture more complex relations in, e.g., images. Long short-term memory networks (LSTMs) [7] correspond to recurrent neural network architectures, able to capture temporal dependencies within data, often used for forecasting tasks.

Traditional machine learning techniques rely on different techniques, that depend on the task to be solved. Random forest [20] is an approach that combines decision trees with bagging and random subspace sampling for classification or regression. The support vector machine (SVM) [21] is a traditional machine learning method relying on maximum-margin hyperplanes for classification, which can also be used for regression.

Reinforcement learning tackles optimization problems, where an agent that interacts with a defined environment must find an optimal solution to the given problem based on a cost function. RL can be combined with deep learning, which is called deep reinforcement learning (DRL). Deep Q Learning [22] is an approach used in deep reinforcement learning where a deep neural network receiving high dimensional data is used to train an RL agent.

3. Methodology

This paper follows the systematic literature review methodology, as specified in the guidelines by Kitchenham et al. [23]. The search process followed is visualized in Fig. 2. The literature search was conducted in the scientific databases IEEEXplore,[1] Scopus,[2] and Web of Science.[3] Scopus and Web of Science were selected, as they provide a large collection of articles from the field of computer and systems sciences. Both Scopus and Web of Science index the publishers Springer, ACM, Wiley, and Taylor & Francis; Due to this, the publishers have not been queried separately. IEEEXplore was included in the search since it is a more specific database focused on technology and engineering. Google Scholar was excluded from our search, as it indexes articles that are not peer-reviewed. Each database was searched with the same parameters, limiting the field to computer science, the publication year between 2002 and 2022 as well as the publication type to journal and conference papers. The publication year was limited until 2022 to guarantee that no newly published papers were indexed anymore, as the literature search was conducted in August 2023.

[1] https://ieeexplore.ieee.org/.
[2] https://www.scopus.com/.
[3] https://www.webofscience.com/.

Table 1
Inclusion criteria.

Criterion	Description
Publication year: 2002–2022	To limit publications between the year of the initial proposal of DT and the last full year
Language: English	To ensure comprehension
Journal or conference paper	To include high quality, peer-reviewed publications
Publication available as a pdf	To review the contents of the publication
Primary study	To restrict the search to original work
Main focus on the use of AI/ML methods within a digital twin	To limit the search to papers focusing on digital twins that employ AI/ML methods
Study depth	Completed research work

Table 2
Characteristics extracted from each relevant paper.

Characteristic	Exemplary values
ML algorithm	LSTM, kNN, CNN, SVM
Algorithm tested	Yes, No
ML task	Classification, Regression, Forecasting
Explainable	Yes, No
Feedback loop	Yes, No
DT represents	Object(s) or process(es) represented
Application domain	Manufacturing, Energy, Healthcare
Task of the AI-DT system	Defect detection, process optimization
Data source of the DT	Theoretical, Synthetic data, Real data
Conceptualization of the DT	Conceptual model, Workflow, Framework
Human in the loop	Yes (including responsibility), No

The choice of search parameters is elaborated in the description of the inclusion criteria in Table 1. The search string used for the query over title, abstract, and keywords is the following:

("Machine Learning" OR "Artificial Intelligence" OR "Deep Learning")
AND ("Digital Twin" OR "Digital Shadow")

The string consists of two parts, which are linked with a logical AND operation. The first part ensures that a keyword related to artificial intelligence is present, which can be machine learning, artificial intelligence, or deep learning. The second part of the query limits the search to papers also containing either digital twin or digital shadow. This ensures that articles terming their solution a digital shadow are included in the analysis. Overall, the query employed is broad, including papers from any domain, that mention artificial intelligence and digital twins.

The search resulted in a total of 2421 papers from the combined search of the databases, which was reduced to 1708 studies after removing duplicates. These articles were filtered based on the inclusion criteria listed in Table 1, with two authors assessing each paper to increase objectivity. The paper inclusion process was done in an iterative way until reaching consensus among the authors. Studies that were considered immature by the authors were excluded, as we are focusing on completed research work. This resulted in a final set of 149 articles, which fulfilled the inclusion criteria and are therefore relevant to the topic of artificial intelligence in digital twins.

This paper investigates the following research questions:

- **RQ1:** How can an artificial intelligence component improve the processing functionality of a digital twin regarding its tasks?
- **RQ2:** Which modeling approaches are used for digital twins employing artificial intelligence in the literature?
- **RQ3:** Are digital twins with artificial intelligence components demonstrating a bidirectional connection between physical and virtual representations?

The first research question (RQ1) is concerned with the overall integration of AI within digital twins regarding the functionality of the twin. The goal of this question is to investigate, which tasks the AI component fulfills and which types of algorithms are commonly implemented to improve the processing functionality. (RQ2) focuses on the modeling approach of digital twins with an AI component, aiming to extract which model-based representations are typically chosen. The last research question (RQ3) examines whether the proposed digital twins implement the characteristic bidirectional connections between virtual and physical systems.

To address the research questions, multiple characteristics were extracted from each of the relevant articles, as shown in Table 2. The characteristics are divided into two categories, one focusing on the AI component and its ML algorithms, the second describing the digital twin and its properties. Table 2 describes each characteristic with example values, while an exhaustive list of values is presented in Section 4.

ML algorithm describes the algorithm or a set of algorithms that were used in a paper. In the case of neural networks, different architectures of the same network type were summarized with the network type, such as LSTM, CNN, or Neural Network. The characteristic *Algorithm tested* can take the values Yes and No, describing whether the paper demonstrated an evaluation of the method, meaning that its performance was tested on a dataset. The *ML task* is divided into the elementary problem types that are

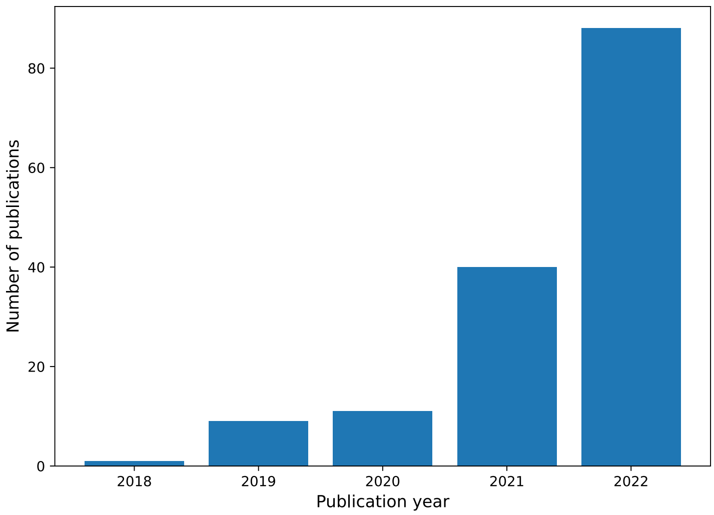

Fig. 3. Histogram of publication years.

encountered in machine learning, such as classification, numeric regression, forecasting of sequential data, or outlier detection. We further categorize each ML algorithm as either *Explainable* or *Non-Explainable* based on its inherent explainability.

The characteristic *Feedback loop* states whether a given research study demonstrates that its proposal for a digital twin has a feedback loop to the physical system. This characteristic takes the value No for the studies that only show a visionary concept of a feedback loop. The object or process that is represented by the digital twin is described with the characteristic *DT represents*, which can also take multiple values when a paper proposes DTs of multiple objects or processes. The *Application domain* characterizes the business domain that the digital twin was applied in. The characteristic *Task of the AI-DT system* shows which task the system was built for, which could be, for example, automatic path planning or anomaly detection. This characteristic is based on the application domain and provides detail from a business-oriented point of view, different from the technically-focused *ML task*. The type of data that the system is using is described by *Data source of the DT*, which can take the values *Theoretical*, where the proposed AI-DT system is not tested with any data, *Synthetic data*, which is artificially generated data, *Real data*, which is historic data stemming from a real system, or *Live data*, which is streamed, real-time data from a real system.

Conceptualization of the DT describes the modeling approach that was used to present the proposed digital twin, and it includes the following approaches: *Schematic model*, a high-level schematic diagram of the DT system, which does not follow any modeling languages; *Workflow*, a textual or graphical description of a temporally ordered procedure for the usage of the model; *Framework*, a combination of a workflow and a schematic model; *System architecture*, an overview of the components of the system showing their interactions and structure; *Conceptual model*, a structured diagram of system concepts with clear relations and cardinalities between them following a modeling language such as Unified Modeling Language (UML), Business Process Modeling Notation (BPMN), or other. To extract this characteristic, the proposed models were classified based on the definitions given in this paragraph. When differing definitions for modeling approaches are used in the papers, we follow the definitions given in this paragraph. Lastly, the characteristic *Human in the loop* states whether a human plays an active role within the AI-DT system, and if there is one, which role the human takes on.

4. Results

This section outlines the results of the literature search, aggregating the extracted characteristics from the final set of 149 relevant articles. For characteristics that can take multiple values for a single study, namely *ML algorithm*, *ML Task*, and *DT represents*, studies were counted once for each of the values, leading to a higher absolute number of occurrences than the total number of relevant papers when evaluating these characteristics. Fig. 3 shows the distribution of publication years of the found studies. The first relevant paper about AI-DT included in our analysis was published in 2018 [24], with the subsequent years showing an increase in popularity. In 2022, 88 relevant papers were published, more than double compared to 2021.

Fig. 4 highlights the imbalance in the distributions of the three *Yes/No* characteristics that we extracted from each paper. 133 papers (89.3%) evaluate their algorithms, while 16 papers only conduct theoretical research, not evaluating their proposals. This shows that research on AI in DTs is typically evaluated in practice and not only proposed on a theoretical basis, while theoretical papers proposing new concepts represent a smaller portion of the overall research.

141 studies (94.6%) do not describe a DT architecture that includes a human as a component of the system. Of the studies that include a human, different roles are taken on: in the domain of healthcare, Tai et al. [25] and Gupta et al. [26] propose a DT that has a doctor in the loop, which is a common use-case where expert knowledge is beneficial for system performance. Latif et al. [27] introduce a framework where a production manager receives recommendations from the DT to improve an assembly process. Shi

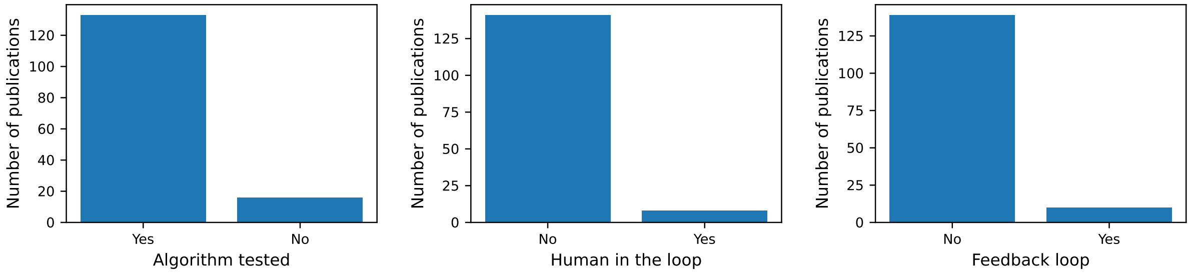

Fig. 4. Imbalanced distribution of studies that tested their algorithms, had a human in the loop, and demonstrated a feedback loop from the virtual system to the physical system.

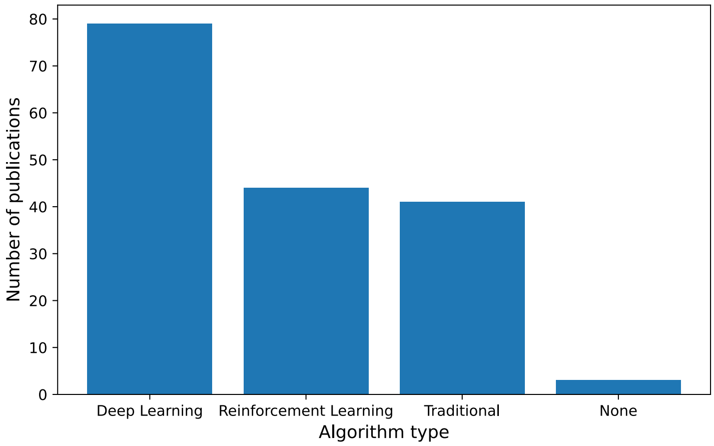

Fig. 5. Histogram of algorithm types.

et al. [28] propose a theoretical model where a human operator collaborates with a robot in a production process. A similar proposal is made by Gallala et al. [29] where an operator works with a robot in a digital twin environment. Pires et al. [30] also introduce a model where an operator interacts with a digital twin of an assembly line with the goal of optimizing productivity. Barricelli et al. [31] propose a fitness digital twin, where a fitness coach, acting as a rule editor, interacts with the DT environment to support the decision-making process. Um et al. [24] utilize smart glasses, integrating the user with a virtual reality environment, serving as the DT.

A total of 138 articles (92.6%) do not clearly demonstrate a feedback loop from the virtual system to the physical system. Overall, only a few papers [32–34] clearly show the effect of their feedback loop, while some papers [35–37] show a visionary feedback loop but do not demonstrate a concrete application of their feedback loop. This highlights that many papers propose a digital twin that does not integrate a virtual-to-physical feedback loop. In these cases, the proposed models do not fulfill a key criterion of a digital twin.

4.1. Machine learning methods used in digital twins

To categorize the ML methods used by the relevant studies, every algorithm was labeled as either *Deep Learning*, *Reinforcement Learning* or *Traditional*. The histogram in Fig. 5 shows the number of publications utilizing each of the techniques. Studies that utilized algorithms from multiple categories were counted once for each category that was used. 3 studies (2.01%) only used preprocessing techniques such as PCA, which were classified as *None*. Deep learning is the most popular machine learning technique, with 79 papers (53.0%) integrating DL with a DT. 44 papers (29.5%) used an AI component based on reinforcement learning, while 41 papers (27.5%) employed traditional machine learning methods. This confirms the popularity of using deep learning approaches to solve complex problems, which has recently been seen in multiple domains [16]. Despite the popularity of deep learning, a considerable number of papers base their research on reinforcement learning and traditional ML methods. Reinforcement learning is designed to tackle different problems than deep learning, which is a possible explanation for their co-existence. Traditional ML methods are often [38–40] used alongside deep learning approaches to compare their performance for the given task of the AI component.

In the 149 studies, a total of 217 ML algorithms were employed, with an average of 1.46 and a median of 1.0 algorithms per article. A total of 36 studies used more than one algorithm, showing that most articles do not make a comparison of different ML

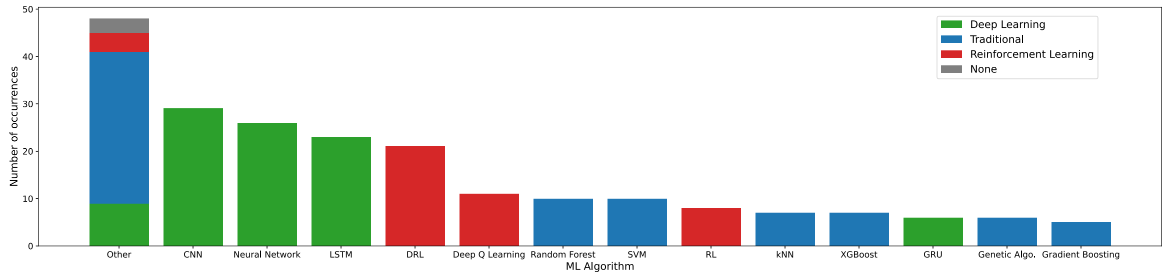

Fig. 6. Histogram of ML algorithms. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

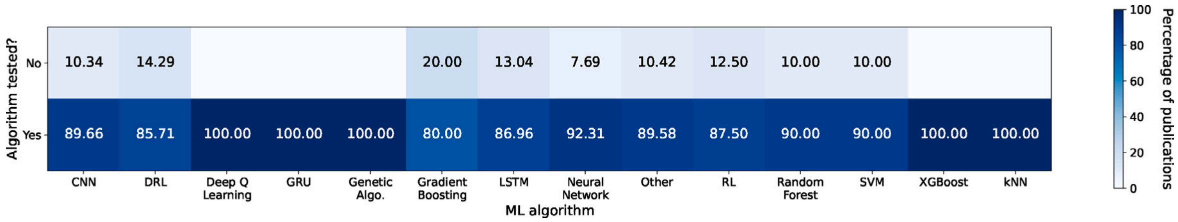

Fig. 7. AI algorithms by testing rate.

solutions. Fig. 6 shows the distribution of ML algorithms, colored by algorithm type. All algorithms that had less than 5 occurrences across all papers were grouped and labeled as *Other* to avoid a long-tailed distribution. Most of the algorithms in *Other* belong to traditional ML methods, with only a small number of other RL and DL methods. The most popular methods are three deep learning methods, namely convolutional neural networks (CNNs) (29 papers, 13.4%), feedforward neural networks (26 papers, 12.0%), and long short-term memory networks (LSTM) (23 papers, 10.6%). These are popular neural network architectures suited for different machine learning tasks. Multiple papers utilize deep reinforcement learning (DRL) (21 papers, 9.68%) and deep Q learning (11 papers, 5.07%). The two most popular traditional ML algorithms are random forest (10 papers, 4.61%), and the support vector machine (SVM) (10 papers, 4.61%). Fig. 7 gives detailed insights on the correlation between AI algorithms and their testing rate on a by-algorithm basis. All algorithms have a testing rate of 80.00% or above, with gradient boosting having the lowest testing rate. Further, Deep Q learning, GRU networks, genetic algorithms, XGBoost, and kNN have a testing rate of 100%. All other algorithms range between 85.71% and 92.31% testing rate. This confirms the statement that the majority of papers test their algorithms. When grouping and normalizing by ML algorithm type, all three groups (DL, RL, traditional) show a high percentage of papers that tested their approaches, between 90.00 and 93.33%. This demonstrates that most researchers proposing an AI-DT system also test their solution and show the performance of the AI component regarding its task with a quantitative analysis. Performance evaluation is task-dependent and individual to each study, therefore, a comparative analysis of algorithmic performance between studies is not conducted.

4.2. Objects represented by digital twins

The extracted characteristic *DT represents* demonstrates which object or process was virtually represented by a digital twin for each publication. As shown in Fig. 8, which does not differentiate between object and process, the majority of papers modeled a single unique type of object (145 studies, 97.3%). Two studies (1.34%) modeled two types of object: Li et al. [41] propose a digital twin system of vehicles and infrastructure for resource optimization. Miao and Zhang [42] use a DT of an unmanned aerial vehicle with a DT of a simulation environment to optimize path planning. Two articles modeled three types of objects: Li et al. [43] model an edge network that consists of mobile terminal users, an unmanned aerial vehicle, and resource devices. Wang et al. [44] model digital twins of humans, cars, and traffic with the goal of vehicle trajectory optimization. Overall, Fig. 8 underlines that most research in the field is focused on applying a digital twin of a single object type, which is not part of a larger system of systems where digital twins of multiple types of objects interact. Aggregating multiple digital twins, with communication between them, would allow for modeling more complex systems, than by modeling a single object.

Fig. 9 shows that 127 publications (85.2%) model an object or multiple objects, while only 22 publications (14.8%) represent a process with their proposed digital twin. None of the publications model both a process and an object. Process modeling is common in the domain of manufacturing, which is highlighted in Fig. 10, which correlates the application domain and the represented object or process. Manufacturing models processes in 39.22% of the publications, while all other domains have either no publications modeling processes or 8.33%, in the case of healthcare and robotics, demonstrating that most domains focus on modeling objects. In healthcare, Chen et al. [45] model COVID-19 disease progression as a process, while in robotics Shi et al. [28] model a human–robot-collaboration process. A possible reason for this is that processes in these domains are typically not formally modeled and are, therefore, difficult to translate into a digital twin. In manufacturing, on the other hand, production processes, such as an assembly process, are well structured and have underlying process models, making a digital twin of such a process more feasible.

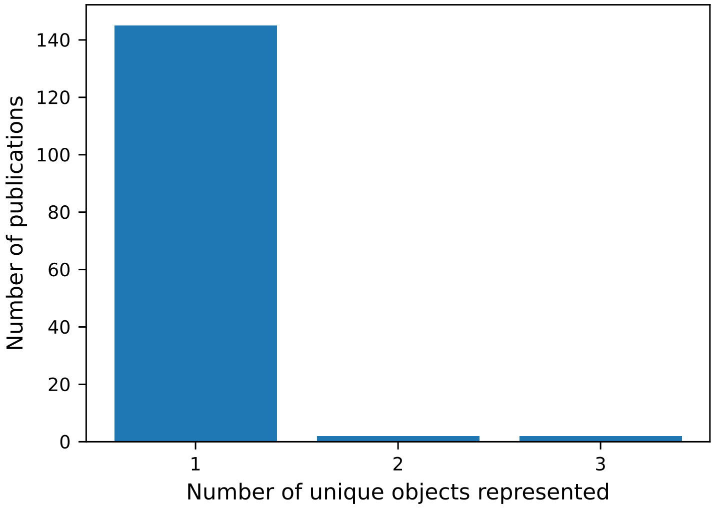

Fig. 8. Number of unique objects represented in each DT.

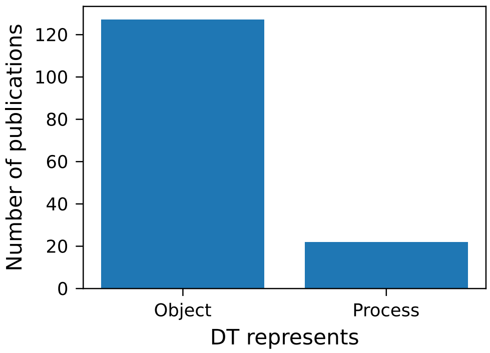

Fig. 9. Number of publications modeling objects and processes.

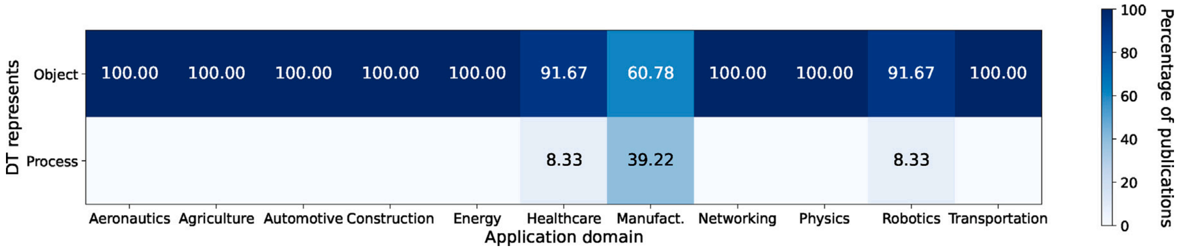

Fig. 10. Publications modeling objects and processes by domain.

4.3. Tasks of AI-DT systems

The tasks of the AI component, and with it, the ML method, can differ depending on the use-case of the DT. Fig. 11 shows the underlying ML tasks that were identified based on the results of the literature search. Papers that pursued multiple tasks within their DT, such as Fahim et al. [5] and Chhetri et al. [46], were counted once for each task, leading to a total of 155 tasks from 149 unique papers. The most common tasks are optimization (51 times, 32.9%), classification (49 times, 31.6%), and regression (34 times, 21.9%). Less common are forecasting (15 times, 9.7%), outlier detection (3 times, 1.9%), and clustering (2 times, 1.3%). Additionally, one article [47], which is not included in the histogram, did not pursue any specific ML task and instead proposed a theoretical architecture for a network digital twin in which multiple different ML tasks can be carried out. It can be argued that clustering and outlier detection are uncommonly seen tasks as they belong to the domain of unsupervised learning, while the other tasks belong to supervised learning and reinforcement learning. This shows that most applications of AI within DT work on

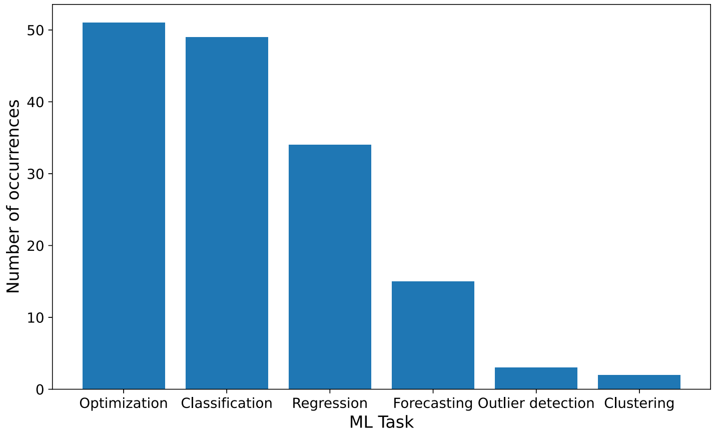

Fig. 11. Distribution of ML tasks.

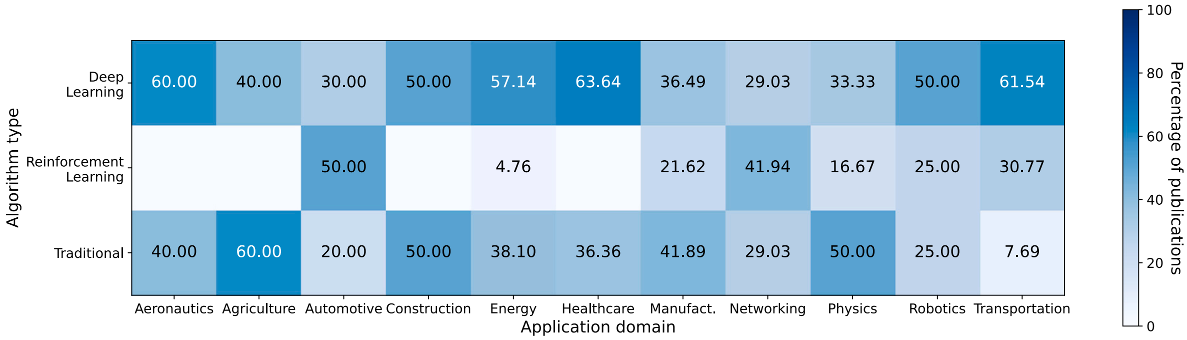

Fig. 12. Types of ML algorithms implemented by application domain.

supervised learning with labeled data or reinforcement learning in an agent-environment scenario. Most optimization tasks are solved with reinforcement learning, while classification, regression, and forecasting belong to supervised learning. The task *classification* was assigned to studies that performed classification on tabular data, images, text, or temporal data, regardless of the input data type, which offers an explanation for the high number of studies working on classification.

The task of the ML component is connected to the overall task of the AI-DT system. This overall task varies depending on the use case, and the domain. In the relevant papers, a total of 98 unique tasks were identified, with the most common being anomaly detection (12 papers, 8.05%), network performance optimization (9 papers, 6.04%), and production process optimization (7 papers, 4.70%). These findings prove that digital twin systems that integrate an AI component have a wide range of application cases in different scenarios.

4.4. Application cases of deep learning, reinforcement learning, and traditional machine learning

Fig. 12 shows the relation between algorithm type and application domain. The heatmap shows percentages that are normalized by the number of papers per application domain. Reinforcement learning methods are primarily applied in a subset of all domains, namely automotive, networking, transportation, robotics, manufacturing, and physics. This demonstrates that different domains are facing different problem types, which require different types of ML algorithms. Aeronautics and agriculture are some of the domains where none of the articles applied reinforcement learning. The applicability of reinforcement learning highly depends on the problem being tackled and the cost of negative outcomes. Both automotive and networking commonly employ reinforcement learning, with over 40% of the papers relying on the technique. This can be explained by the simulation capabilities of a digital twin, which allow for solving complex optimization problems such as automatic path planning for cars or network resource optimization. In agriculture, manufacturing, and physics, traditional ML approaches are the most dominant solutions, while deep learning is more common for most other domains.

Fig. 13 visualizes the differences between deep learning, reinforcement learning, and traditional approaches, regarding the data source of the DT. The graph shows percentages of publications normalized by the ML algorithm type. Both traditional approaches and DL show a similar distribution, with data stemming from a real system being the most common data source, followed by synthetic data, and a small percentage of theoretical approaches and approaches relying on live data. RL shows different results, with no

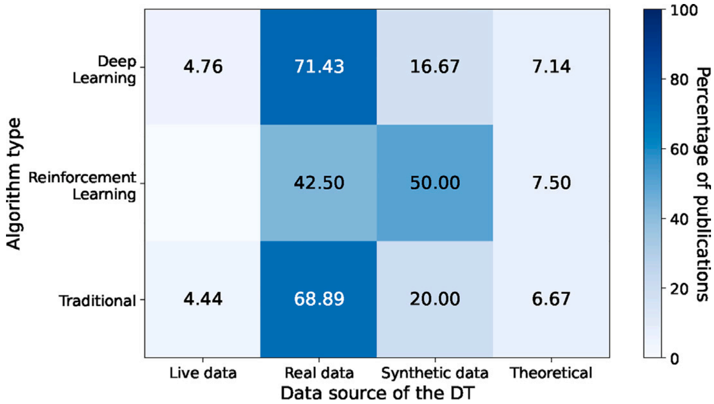

Fig. 13. Types of AI algorithms and data source of the DT.

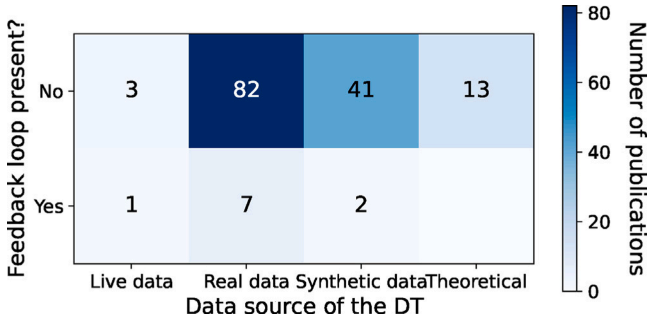

Fig. 14. Relation of the data source of the DT and presence of a feedback loop.

papers utilizing live data, a lower percentage of papers using real data, and a higher percentage using synthetic data. This can be attributed to the fact that RL does not rely on datasets, but requires an environment to operate, which is often synthetically created, rather than using the specification of a real system. Overall, RL relies on less real-world data than DL and traditional approaches, showing that the application cases of RL in digital twins are less mature.

Of the 149 relevant papers, only one paper [48] proposes a digital twin operating with live data that shows both a physical-to-virtual and a virtual-to-physical connection, demonstrating a clear feedback loop. The authors propose a digital twin for iron reverse flotation, a chemical production process. They train their AI component based on historical data and integrate their model with real-time data from a live system, observing changes in productivity after switching to the digital twin-based system. Their feedback loop dynamically adapts the dosage used in the production system, optimizing real-time productivity. In total, 7.38% of the studies (11 papers) present a feedback loop from the digital twin to the physical system. Additionally, only 2.68% (4 papers) of the analyzed publications work with live data stemming from a real system, which allows for online training of an AI component and decision-making based on real-time data. Recalling the criteria that define a digital twin, as given by Grieves [14], real-time data, and a bidirectional data flow between virtual and physical systems are essential parts of a digital twin. Fig. 14 clearly demonstrates a gap in current research where solutions are termed "digital twin", but do not fulfill the requirements for a digital twin.

4.5. Conceptualization approaches for digital twins

Every paper included in our analysis was categorized by the conceptualization approach of the digital twin proposed. When no clear model was shown, this category takes the value *None*, which was the case for 3 studies (2.01%), as shown in Fig. 15. 53 articles (35.6%) show a schematic model of a digital twin, which is a model on a high abstraction level. The second most common approach is a framework, which was used by 51 papers (34.2%). System architectures, which are more detailed models, were proposed in 30 publications (20.1%), while 12 papers (8.05%) conceptualized their DT with a workflow. In the criteria we used, *conceptual model* was another possibility for conceptualization approaches, however, none of the found papers showed a conceptual model. Overall, it can be stated that most of the models shown in the literature on AI-DT systems are immature, providing a shallow overview of the system, often only schematically describing the proposed solution.

Fig. 16 presents the relation between algorithm types and the conceptualization approach of the digital twin. The plot is normalized by the conceptualization approach to show the relative differences between them. Deep learning is the most common approach, with between 50.85% and 53.33% for all types of conceptualization except for system architectures. System architectures

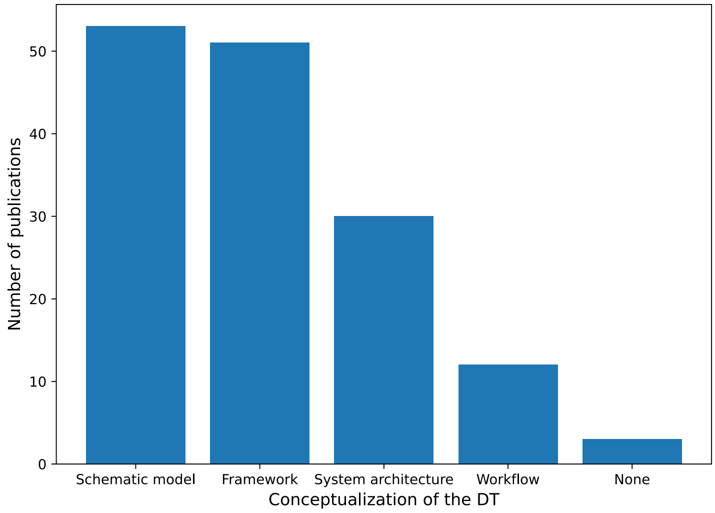

Fig. 15. Conceptualization approaches of the DTs.

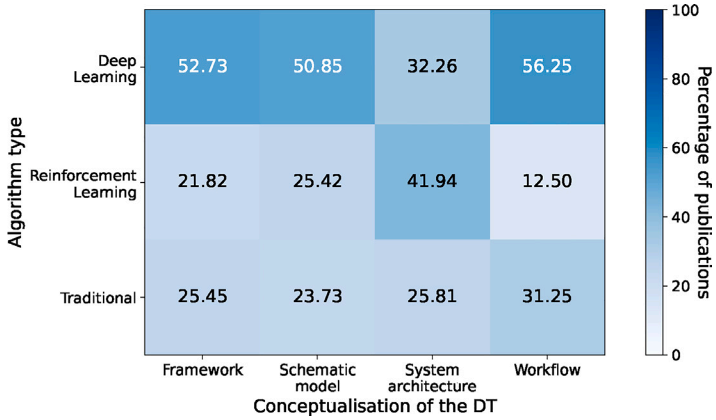

Fig. 16. Conceptualization of digital twins relating to types of AI algorithms.

show the highest percentage of studies utilizing reinforcement learning, with 43.33%. Workflows show a significantly lower percentage of papers using reinforcement learning, with only 13.33%. A possible reason is that reinforcement learning requires a clear specification of an environment, an actor, and a reward system, which aligns with the requirements to design a system architecture, while a workflow can be less specific. Workflows show considerably more usage of traditional ML algorithms (33.33%) than the other approaches of conceptualization (23.33%–25.45%), which can be attributed to the fact that workflows have a relatively lower number of papers implementing reinforcement learning.

4.6. Explainability of an AI-DT system

Model explainability in the context of AI-DT systems refers to the self-explanatory nature of a digital twin model, where more detailed models are more explainable, and higher-level models are less explainable. As seen in Fig. 15, most papers use high-level modeling approaches, that only provide little explainability. This aligns with the results regarding algorithmic explainability, showing that overall, most AI-DT systems do not consider explainability, either by model explainability or algorithmic explainability.

Algorithmic explainability refers to an ML algorithm being a white-box algorithm, where predictions can be understood by a domain expert, while non-explainable algorithms, such as neural networks, are black-box solutions. Some algorithms can potentially be explainable, depending on the use case. Fig. 17 provides a histogram of the number of studies that have used explainable algorithms. We classified the algorithms as either *Explainable* or *Non-explainable*. Examples of explainable algorithms are decision

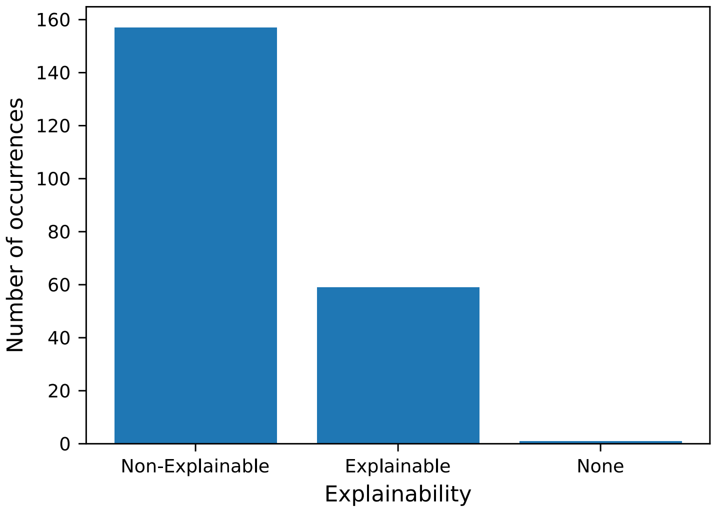

Fig. 17. Histogram of algorithm explainability across the proposed DTs.

trees or linear regression models, while non-explainable networks are for convolutional neural networks or recurrent neural networks. 59 (27.19%) of the used algorithms are explainable, while 157 (72.4%) are less explainable. This demonstrates that algorithmic explainability is an often overlooked factor which is crucial for a human operator to understand the predictions of the AI component.

5. Discussion

Low number of studies clearly demonstrating a feedback loop and using live data

We identified multiple gaps in the literature at the intersection of artificial intelligence and digital twins based on the results of this literature review. Firstly, most papers in the field do not define nor demonstrate a feedback loop, i.e. the virtual-to-physical connection, that utilizes the output of the digital twin. This is one of the main criteria that define a digital twin [14] since it allows the virtual system to not only simulate but also influence the real-world system. Additionally, most studies that were included in our search did not base their system on real-time data, which is another key capability of a digital twin. This is aligned with the findings of a recent systematic literature review by Wooley et al. [49], who investigate the difference between simulations and digital twins, while not specifically searching only for studies in the field of AI-DT. The authors state that an issue common to many studies is that traditional simulations are termed digital twins, despite not defining or demonstrating real-time data synchronization between virtual and physical systems. We made similar observations, with many articles showing a feedback loop in schematic diagrams, but not providing details on it. In our results, only one paper presents a digital twin with an AI component, that operates with live data and clearly demonstrates a feedback loop [48].

Low number of proposed digital twins with a human in the loop

A small subset of the analyzed papers integrated a human as a key part of the AI-DT system. It has been shown that human-in-the-loop machine learning can improve both model performance and explainability [50], two desirable traits for an AI-DT system. This is a possibility for future research work to expand upon, by either improving past research or proposing new model architectures that integrate a human. In the set of papers found in the literature search where a human interacted with the DT system, the human took different roles, such as operator, doctor, rule editor, or production manager. The current state of the art lacks a clear definition of possible roles for humans within an AI-DT system, which is applicable across different domains, providing human knowledge at different points in the system. Human-in-the-loop systems are key for Industry 5.0, which is often characterized as human-centric in the literature [51].

Lack of digital twins modeling multiple objects

The majority of papers included in our review propose a digital twin for a single type of object. We found a research gap, with only a small minority of analyzed studies combining digital twins of multiple object types. The next logical step is connecting multiple digital twins in a system of systems-based approach. Tao et al. [52] have suggested hierarchical levels of digital twins in manufacturing, where a system of systems is the highest level of abstraction. On top of the requirements for building digital twins, interoperability, data synchronicity, and communication [53] between the DTs need to be considered to design a system of systems with multiple DTs.

Processes are modeled almost exclusively in manufacturing

The digital twins in the surveyed papers typically model objects, with a smaller portion modeling processes. We found that processes are almost exclusively modeled in the domain of manufacturing. This is related to the emergence of Industry 4.0, which has accelerated the adoption of digital twins in the manufacturing sector. By modeling processes in other application domains with an AI-DT system, future work can address this gap. Some domains, however, may naturally be based on fewer processes and more objects and are, therefore, more inclined to replicate these objects. A prime example of this is construction, where the objects of interest are tunnels [54], buildings [55], or wind turbines [36]. Modeling the process of constructing a structure is less common, which also reflects in our data, although digital twins of the construction phase, that do not employ ML techniques, have been proposed [56].

Reinforcement learning is not used in some application domains

Reinforcement learning in digital twins has seen considerable attention in some domains, such as automotive, networking, and transportation, where the digital twin is used as a simulation environment for the RL agent. However, in other domains, namely aeronautics, agriculture, construction, and healthcare, none of the studies analyzed in our review make use of reinforcement learning. Due to its nature, reinforcement learning is suited to solve optimization problems that have different problem settings from classification or regression tasks. A digital twin can support a reinforcement learning solution by providing an accurate, low-cost simulation environment, while the RL benefits the DT by learning to solve an optimization problem.

Reinforcement learning-based studies do not work with live data

We found that none of the papers applying reinforcement learning within digital twins use real-time data for their digital twin. Additionally, past work based on RL methods uses synthetic data more often than real data. This confirms that digital twins using RL are often proposed in a lab setting with synthetic data. Considering the definitions of digital twin, digital shadow, and digital model given by Jones et al. [1], a digital twin without live data should instead be termed a digital model. A direction for future research is to investigate how well digital twins using RL integrate with a real-world, real-time data setting compared to the synthetic lab setting.

Most modeling approaches are high-level and do not follow modeling languages

Digital twins in the field of AI-DT are often modeled with schematic models, without following conventional modeling approaches. The majority of proposed models display a high-level overview of the digital twin, i.e., providing a shallow level of detail. More detailed models, such as a system architecture, are shown less commonly, while none of the found papers describe a conceptual model following a modeling language. This showcases the need for more detailed modeling approaches in the AI-DT community, by moving away from schematic models to detailed, in-depth models of the proposed system, which would contribute to a better understanding of the AI-DT system and the ability to increase the efficiency of its design, development, and maintenance.

Lack of model explainability and algorithmic explainability

Most of the articles taken into account in this review do not focus on explainability, both on a model level regarding the digital twin and on the algorithmic level of the AI component. Since digital twins are systems that are designed for real-world scenarios, making both explainability on a model level and on an algorithm level highly desirable properties. To achieve model-level explainability, more detailed, in-depth modeling approaches can be used to describe the proposed digital twins. As deep learning models are inherently not explainable, using post-hoc explainability methods in combination with them can provide algorithmic explanations while also maintaining predictive performance. Alternatively, white-box algorithms, such as decision trees, can be integrated with a DT to achieve the same goal. However, the explanations from post-hoc methods differ from white-box explanations, often adding uncertainty [57].

Variety of tasks and machine learning algorithms used in digital twins

In the analyzed studies, deep learning, reinforcement learning, and traditional machine learning algorithms see consistent usage. This highlights that the problems tackled by digital twins are diverse and require the use of different ML algorithms. From our analysis, it becomes clear that most digital twins are proposed for a unique type of problem, while only a few problems are common across multiple studies. The diversity of problems that digital twins are designed for is aligned with the fact that the AI component in digital twins pursues different tasks, such as optimization, classification, or regression. Overall, AI in digital twins typically works with supervised learning or reinforcement learning, and only rarely with unsupervised learning. A possible reason for this is that digital twins mainly work with labeled data while being less like to perform exploratory tasks on unlabeled data. Additionally, most publications in the field test their algorithms' performance, providing concrete evidence of their predictive power.

RQ1: How can an artificial intelligence component improve the processing functionality of a digital twin regarding its tasks?

In this review, we found that AI components can fulfill a variety of tasks within a digital twin. On the algorithmic level, tasks such as optimization, classification, and regression are commonly seen. Digital twins with an AI component can tackle a broader range of problems, which require predictive functionality, with most proposed twins focusing on a unique problem. An example of a problem that can be tackled by integrating AI with the DT is the forecast of temperatures in a building, allowing adaptive control of heating and ventilation. Depending on the task of the DT, the AI component fulfills its role as a predictive algorithm, relying on the data of the DT to make predictions. Further, without an AI component, certain tasks, such as forecasting of temporal data streams, could not be tackled by a DT on its own, showing that the integration of AI with DT opens the possibility to approach new problem types. It became clear that in the literature, a number of different ML approaches are used for the AI component. This demonstrates that different ML approaches are suited for different tasks, and due to the given variety of tasks, a similar variety in algorithms can be observed.

RQ2: Which modeling approaches are used for digital twins employing artificial intelligence in the literature?

AI-DT systems are commonly modeled with high-level schematic models. Most research work in the field focuses on the implementation and evaluation of the AI component, with little focus being given to the modeling part of the DT. Conceptual models are not present in the literature analyzed in this review, which proves that the field is in need of more detailed, in-depth models for digital twins integrating an AI component. Overall, about 20% of the studies provide a system architecture for the proposed digital twin, which is modeling the DT at a lower level, giving more attention to detail and more clearly specifying the functionality of the system.

RQ3: Are digital twins with artificial intelligence components demonstrating a bidirectional connection between physical and virtual representations?

The bidirectional connection between physical and virtual representation consists of two parts: Firstly, the physical-to-virtual connection, which is the data stream that supplies the digital twin; Secondly, the virtual-to-physical connection, which is the feedback loop from the DT to the physical system. Most publications in the field work with historical data stemming from a real system, while only a small fraction works with live data, providing an automated data flow between physical and virtual systems. This illustrates that, although most proposed DTs are based on real data, the physical-to-virtual connection still follows a manual procedure. The virtual-to-physical connection is demonstrated by a small number of papers, where a feedback loop to the real system is shown. Although articles commonly show a visionary feedback loop in schematic models, for most articles, this feedback loop is not implemented in practice, and a manual information flow between digital twin and real-world system is necessary. In summary, the majority of digital twins presented in the analyzed papers either lacked an implementation of the physical-to-virtual connection or did not demonstrate a virtual-to-physical connection.

6. Conclusion

In this study, we reviewed the state of the art at the intersection of artificial intelligence and digital twins, analyzing 149 relevant articles that propose a digital twin integrating an AI component. The DTs found in the literature model a variety of objects and processes from different application domains. By utilizing an AI component, DTs can tackle a broad range of problems, such as robot collision avoidance or adaptive production scheduling, which are often unique to specific domains. We found that deep learning, reinforcement learning, and traditional ML approaches are applied to overcome these problems. Further, the focus of many publications in the field is on the evaluation of the AI component, which is demonstrated in the majority of papers. However, the digital twins are often modeled on a high level of abstraction, with schematic models providing only a shallow level of detail. Only few studies introduce a DT that relies on live data or clearly demonstrates a virtual-to-physical connection. This highlights the need for future work to focus on the automated bidirectional information flow between physical and virtual systems to make the proposed digital twins suitable for usage with a real-world system.

CRediT authorship contribution statement

Tim Kreuzer: Conceptualization, Data curation, Formal analysis, Investigation, Software, Visualization, Writing – original draft. **Panagiotis Papapetrou:** Conceptualization, Supervision, Validation, Writing – review & editing. **Jelena Zdravkovic:** Conceptualization, Funding acquisition, Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors thank Atrium Ljungberg AB for their insights on challenges in the smart building industry and the resulting need for AI-DT systems.

Appendix. List of analyzed studies

See Table A.3.

Table A.3
List of analyzed studies.

Author	Year	Title
AlZyoud et al.	2022	Towards a Machine Learning-Based Digital Twin for Non-Invasive Human Bio-Signal Fusion
Alamin et al.	2022	A Machine Learning-based Digital Twin for Electric Vehicle Battery Modeling
Alexopoulos et al.	2020	Digital twin-driven supervised machine learning for the development of artificial intelligence applications in manufacturing
Allison et al.	2022	Digital Twin-Enhanced Methodology for Training Edge-Based Models for Cyber Security Applications
Angin et al.	2020	Agrilora: A digital twin framework for smart agriculture
Bandara et al.	2022	Modeling a Digital Twin to Predict Battery Deterioration with Lower Prediction Error in Smart Devices: From the Internet of Things Sensor Devices to Self-Driving Cars
Bansal et al.	2019	Ant colony optimization algorithm for industrial robot programming in a digital twin
Barricelli et al.	2020	Human Digital Twin for Fitness Management
Barriga et al.	2022	Advanced data modeling for industrial drying machine energy optimization
Benzon et al.	2022	An Operational Image-Based Digital Twin for Large-Scale Structures
Berghe	2021	A processing architecture for real-time predictive smart city digital twins
Boulfani et al.	2020	Anomaly detection for aircraft electrical generator using machine learning in a functional data framework
Chakrabarti et al.	2021	Efficient Modeling of Digital Shadows for Production Processes: A Case Study for Quality Prediction in High Pressure Die Casting Processes
Chen et al.	2022	Artificial intelligence enabled Digital Twins for training autonomous cars
Chen et al.	2022	Digital Twin based Train Delay Prediction System: Design and Realization
Chen et al.	2022	Digital twins to fight against COVID-19 pandemic
Chhetri et al.	2019	QUILT: Quality Inference from Living Digital Twins in IoT-Enabled Manufacturing Systems
Chiurco et al.	2022	Data Modeling and ML Practice for Enabling Intelligent Digital Twins in Adaptive Production Planning and Control
Chukkapalli et al.	2021	Cyber-Physical System Security Surveillance using Knowledge Graph based Digital Twins - A Smart Farming Usecase
Cronrath et al.	2019	Enhancing Digital Twins through Reinforcement Learning
Damit et al.	2021	Digital-twin-assisted Software-defined PON: A Cognition-driven Framework for Energy Conservation
Dang et al.	2022	Cloud-Based Digital Twinning for Structural Health Monitoring Using Deep Learning
Dehghanimo. et al.	2021	Simulation-Optimization of Digital Twin
Deng et al.	2021	A Digital Twin Approach for Self-optimization of Mobile Networks
Dong et al.	2019	Deep Learning for Hybrid 5G Services in Mobile Edge Computing Systems: Learn from a Digital Twin
Du et al.	2021	Digital Twin Based Trajectory Prediction for Platoons of Connected Intelligent Vehicles
Du et al.	2022	Segmentation, Detection, and Tracking of Stem Cell Image by Digital Twins and Lightweight Deep Learning
Elayan et al.	2021	Digital Twin for Intelligent Context-Aware IoT Healthcare Systems
Fahim et al.	2022	Machine Learning-Based Digital Twin for Predictive Modeling in Wind Turbines
Fan et al.	2021	Digital Twin Empowered Mobile Edge Computing for Intelligent Vehicular Lane-Changing
FerriolGalmes et al.	2022	Building a Digital Twin for network optimization using Graph Neural Networks
Fraser et al.	2021	Enhancing the Security of Unmanned Aerial Systems using Digital-Twin Technology and Intrusion Detection
Gallala et al.	2022	Digital Twin for Human–Robot Interactions by Means of Industry 4.0 Enabling Technologies
Gao et al.	2022	Machine Learning and Digital Twin-Based Path Planning for AGVs at Automated Container Terminals
Ghandar et al.	2021	A Decision Support System for Urban Agriculture Using Digital Twin: A Case Study With Aquaponics
Goodwin et al.	2022	Real-time digital twin-based optimization with predictive simulation learning
Guerra et al.	2019	Digital Twin-Based Optimization for Ultraprecision Motion Systems With Backlash and Friction
Guo et al.	2021	Fault diagnosis of intelligent production line based on digital twin and improved random forest
Gupta et al.	2021	Hierarchical Federated Learning based Anomaly Detection using Digital Twins for Smart Healthcare
Hu et al.	2022	A grasps-generation-and-selection convolutional neural network for a digital twin of intelligent robotic grasping
Huang et al.	2022	AI-Driven Digital Process Twin via Networked Digital Process Chain
Huang et al.	2022	Network Selection and QoS Management Algorithm for 5G Converged Shipbuilding Network Based on Digital Twin
Huang et al.	2022	Personalized QoE Enhancement for Adaptive Video Streaming: A Digital Twin-Assisted Scheme
Hui et al.	2021	Time or Reward: Digital-twin Enabled Personalized Vehicle Path Planning
Jafari et al.	2022	Prediction of the Battery State Using the Digital Twin Framework Based on the Battery Management System
Kannari et al.	2022	Energy-data-related digital twin for office building and data centre complex
Kharlamova et al.	2022	A Digital Twin of Battery Energy Storage Systems Providing Frequency Regulation
Latif et al.	2020	A Case Study of Digital Twin for a Manufacturing Process Involving Human Interactions
Li	2022	Fault Prediction and Diagnosis System for Large-diameter Auger Rigs Based on Digital Twin and BP Neural Network
Li et al.	2022	A flexible manufacturing assembly system with deep reinforcement learning
Li et al.	2023[a]	An AR-assisted Deep Reinforcement Learning-based approach towards mutual-cognitive safe human–robot interaction
Li et al.	2022	Digital Twin Assisted Task Offloading for Aerial Edge Computing and Networks

(continued on next page)

Table A.3 (*continued*).

Author	Year	Title
Li et al.	2021	Digital Twin Enhanced Assembly Based on Deep Reinforcement Learning
Li et al.	2022	Digital Twin-Driven Computing Resource Management for Vehicular Networks
Li et al.	2022	Digital Twin-Driven Task Replanning Method for Robot-Environment Physical Interaction
Li et al.	2022	The Digital Twin Model of Chemical Production Systems in Smart Factories: A Case Study
Lin et al.	2021	Evolutionary digital twin: A new approach for intelligent industrial product development
Liu et al.	2022	A digital twin-based sim-to-real transfer for deep reinforcement learning-enabled industrial robot grasping
Liu et al.	2022	Adaptive reconstruction of digital twins for machining systems: A transfer learning approach
Liu et al.	2021	CNC Machine Tool Fault Diagnosis Integrated Rescheduling Approach Supported by Digital Twin-Driven Interaction and Cooperation Framework
Liu et al.	2021	Design of Photovoltaic Power Station Intelligent Operation and Maintenance System Based on Digital Twin
Liu et al.	2022	Digital Twins-Based Impact Response Prediction of Prestressed Steel Structure
Livera et al.	2022	Intelligent Cloud-Based Monitoring and Control Digital Twin for Photovoltaic Power Plants
Lu et al.	2021	Low-Latency Federated Learning and Blockchain for Edge Association in Digital Twin Empowered 6G Networks
Luo et al.	2021	A Digital Twin-Driven Methodology for Material Resource Planning Under Uncertainties
Lv et al.	2022	Traffic Safety Detection System by Digital Twins and Virtual Reality Technology
Ma et al.	2022	Using Deep Reinforcement Learning for Zero Defect Smart Forging
Maia et al.	2022	Holistic Security and Safety for Factories of the Future
Manocha et al.	2023	Digital Twin-assisted Blockchain-inspired irregular event analysis for eldercare
Martinez et al.	2021	Automation pyramid as constructor for a complete digital twin, case study: A didactic manufacturing system
MartinezVelazquez et al.	2019	Cardio Twin: A Digital Twin of the human heart running on the edge
Matulis et al.	2021	A robot arm digital twin utilising reinforcement learning
Melesse et al.	2022	Machine Learning-Based Digital Twin for Monitoring Fruit Quality Evolution
Miao et al.	2022	UAV Visual Navigation System based on Digital Twin
Min et al.	2019	Machine Learning based Digital Twin Framework for Production Optimization in Petrochemical Industry
Molinaro et al.	2021	Embedding data analytics and CFD into the digital twin concept
Mortlock et al.	2022	Graph Learning for Cognitive Digital Twins in Manufacturing Systems
Mozo et al.	2022	B5GEMINI: AI-Driven Network Digital Twin
MuellerZhang et al.	2020	Dynamic Process Planning using Digital Twins and Reinforcement Learning
Mulinka et al.	2022	Optimizing a Digital Twin for Fault Diagnosis in Grid Connected Inverters - A Bayesian Approach
MullerZhang et al.	2022	A Digital Twin-based Approach Performing Integrated Process Planning and Scheduling for Service-based Production
Neuburger et al.	2022	Coupled Finite-Element-Method-Simulations for Real-Time-Process Monitoring in Metal Forming Digital-Twins
Okegbile et al.	2022	Edge-assisted human-to-virtual twin connectivity scheme for human digital twin frameworks
Pang et al.	2021	Collaborative City Digital Twin for the COVID-19 Pandemic: A Federated Learning Solution
Park et al.	2022	Information fusion and systematic logic library-generation methods for self-configuration of autonomous digital twin
Piltan et al.	2022	Strict-Feedback Backstepping Digital Twin and Machine Learning Solution in AE Signals for Bearing Crack Identification
Piper et al.	2022	Digital Twins for Smart Cities: Case Study and Visualisation via Mixed Reality
Pires et al.	2021	Recommendation System Using Reinforcement Learning for What-If Simulation in Digital Twin
Poore et al.	2022	Multi-Physics and Artificial Intelligence Models for Digital Twin Implementations of Residential Electric Loads
Ren et al.	2022	Machine-Learning-Driven Digital Twin for Lifecycle Management of Complex Equipment
Ren et al.	2021	Strengthening Digital Twin Applications based on Machine Learning for Complex Equipment
Ritto et al.	2021	Digital twin, physics-based model, and machine learning applied to damage detection in structures
Salim et al.	2022	A Blockchain-Enabled Secure Digital Twin Framework for Early Botnet Detection in IIoT Environment
Santos et al.	2021	Decision-making in a fast fashion company in the Industry 4.0 era: a Digital Twin proposal to support operational planning
SerranoRuiz et al.	2021	Smart Master Production Schedule for the Supply Chain: A Conceptual Framework
Shen et al.	2022	Deep Reinforcement Learning for Flocking Motion of Multi-UAV Systems: Learn From a Digital Twin
Shi et al.	2022	A Cognitive Digital Twins Framework for Human-Robot Collaboration
Shi et al.	2022	Visual Human Localization and Safety Monitoring in a Digital Twin of Workspace
Snijders et al.	2020	Machine Learning for Digital Twins to Predict Responsiveness of Cyber-Physical Energy Systems
Song et al.	2021	Adaptive Federated Learning for Digital Twin Driven Industrial Internet of Things
Sun et al.	2021	Digital twins in human understanding: a deep learning-based method to recognize personality traits
Sun et al.	2022	Dynamic Digital Twin and Federated Learning with Incentives for Air-Ground Networks

(*continued on next page*)

Table A.3 (*continued*).

Author	Year	Title
Sun et al.	2020	Reducing Offloading Latency for Digital Twin Edge Networks in 6G
Tai et al.	2022	Digital-Twin-Enabled IoMT System for Surgical Simulation Using rAC-GAN
Tancredi et al.	2022	Integration of Digital Twin, Machine-Learning and Industry 4.0 Tools for Anomaly Detection: An Application to a Food Plant
Tang et al.	2022	DT-EEC: A Digital Twin-assisted End-Edge-Cloud Collaboration Architecture for Industrial Internet
To et al.	2021	Drone-Based AI and 3D Reconstruction for Digital Twin Augmentation
Tu et al.	2022	Digital Twins-Based Automated Pilot for Energy-Efficiency Assessment of Intelligent Transportation Infrastructure
Um et al.	2018	Modular augmented reality platform for smart operator in production environment
Varghese et al.	2022	Digital Twin-based Intrusion Detection for Industrial Control Systems
Wang et al.	2022	A Graph Neural Network-Based Digital Twin for Network Slicing Management
Wang et al.	2022	Adaptive Optimization Method in Digital Twin Conveyor Systems via Range-Inspection Control
Wang et al.	2020	Deep learning-empowered digital twin for visualized weld joint growth monitoring and penetration control
Wang et al.	2021	Digital Twin for Human-Robot Interactive Welding and Welder Behavior Analysis
Wang et al.	2022	Mobility Digital Twin: Concept, Architecture, Case Study, and Future Challenges
Wang et al.	2022	Real-Time Analysis of Multiple Root Causes for Anomalies Assisted by Digital Twin in NFV Environment
Wehner et al.	2022	Explainable Online Lane Change Predictions on Digital Twin with Layer Normalized LSTM and Layer-wise Relevance Propagation
Wu et al.	2021	A Framework of Dynamic Data Driven Digital Twin for Complex Engineering Products: the Example of Aircraft Engine Health Management
Wu et al.	2021	Digital Twin-enabled Reinforcement Learning for End-to-end Autonomous Driving
Wu et al.	2022	Digital twins and artificial intelligence in transportation infrastructure: Classification, application, and future research directions
Xie et al.	2019	A Neural Ordinary Differential Equations Based Approach for Demand Forecasting within Power Grid Digital Twins
Xie et al.	2022	Digital Twin Enabled Dual-System Reinforcement Learning Method
Xing et al.	2022	Multi-energy Simulation and Optimal Scheduling Strategy Based on Digital Twin
Xiong et al.	2021	Digital twin–driven aero-engine intelligent predictive maintenance
Xu et al.	2019	A Digital-Twin-Assisted Fault Diagnosis Using Deep Transfer Learning
Xu et al.	2021	Digital Twin-based Anomaly Detection in Cyber-physical Systems
Xu et al.	2022	Dynamic Scheduling of Crane by Embedding Deep Reinforcement Learning into a Digital Twin Framework
Xu et al.	2022	Service Offloading With Deep Q-Network for Digital Twinning-Empowered Internet of Vehicles in Edge Computing
Yu et al.	2023	Edge intelligence-driven digital twin of CNC system: Architecture and deployment
Yuan et al.	2022	Digital Twin-Driven Vehicular Task Offloading and IRS Configuration in the Internet of Vehicles
Zhang et al.	2020	A Product Quality Monitor Model With the Digital Twin Model and the Stacked Auto Encoder
Zhang et al.	2022	A digital twin dosing system for iron reverse flotation
Zhang et al.	2022	Adaptive Digital Twin and Multiagent Deep Reinforcement Learning for Vehicular Edge Computing and Networks
Zhang et al.	2020	Deep learning-enabled intelligent process planning for digital twin manufacturing cell
Zhang et al.	2022	Digital Twin Empowered Content Caching in Social-Aware Vehicular Edge Networks
Zhang et al.	2022	Real Time Object Detection in Digital Twin with Point-Cloud Perception for a Robotic Manufacturing Station
Zhang et al.	2022	Secure medical digital twin via human-centric interaction and cyber vulnerability resilience
Zhang et al.	2022	Smart DC: An AI and Digital Twin-based Energy-Saving Solution for Data Centers
Zhao et al.	2022	An Established Theory of Digital Twin Model for Tunnel Construction Safety Assessment
Zhao et al.	2022	Digital Twin-Driven Estimation of State of Charge for Li-ion Battery
Zhao et al.	2022	Digital twin for rapid damage detection of a fixed net panel in the sea
Zheng et al.	2022	Digital Twin Empowered Heterogeneous Network Selection in Vehicular Networks With Knowledge Transfer
Zheng et al.	2021	Digital twin-trained deep convolutional neural networks for fringe analysis
Zhou et al.	2022	Digital Twin-Empowered Network Planning for Multi-Tier Computing
Zhou et al.	2021	Intelligent Ironmaking Optimization Service on a Cloud Computing Platform by Digital Twin
Zhou et al.	2022	Intelligent Small Object Detection for Digital Twin in Smart Manufacturing With Industrial Cyber-Physical Systems
Zhou et al.	2022	Secure and Latency-Aware Digital Twin Assisted Resource Scheduling for 5G Edge Computing-Empowered Distribution Grids
Zohdi	2021	A Digital-Twin and Machine-Learning Framework for Ventilation System Optimization for Capturing Infectious Disease Respiratory Emissions
Zohdi	2022	A digital-twin and machine-learning framework for precise heat and energy management of data-centers

[a] A small number of studies published in 2023 are indexed with the publication year 2022.

References

[1] D. Jones, C. Snider, A. Nassehi, J. Yon, B. Hicks, Characterising the digital twin: A systematic literature review, CIRP J. Manuf. Sci. Technol. 29 (2020) 36–52, http://dx.doi.org/10.1016/j.cirpj.2020.02.002.

[2] M. Helmstaedter, K.L. Briggman, S.C. Turaga, V. Jain, H.S. Seung, W. Denk, Connectomic reconstruction of the inner plexiform layer in the mouse retina, Nature 500 (7461) (2013) 168–174, http://dx.doi.org/10.1038/nature12346.

[3] D. Silver, A. Huang, C.J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, et al., Mastering the game of go with deep neural networks and tree search, Nature 529 (7587) (2016) 484–489, http://dx.doi.org/10.1038/nature16961.

[4] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, et al., Highly accurate protein structure prediction with AlphaFold, Nature 596 (7873) (2021) 583–589, http://dx.doi.org/10.1038/s41586-021-03819-2.

[5] M. Fahim, V. Sharma, T.-V. Cao, B. Canberk, T.Q. Duong, Machine learning-based digital twin for predictive modeling in wind turbines, IEEE Access 10 (2022) 14184–14194, http://dx.doi.org/10.1109/ACCESS.2022.3147602.

[6] M. Xiong, H. Wang, Q. Fu, Y. Xu, Digital twin–driven aero-engine intelligent predictive maintenance, Int. J. Adv. Manuf. Technol. 114 (11–12) (2021) 3751–3761, http://dx.doi.org/10.1007/s00170-021-06976-w.

[7] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780, http://dx.doi.org/10.1162/neco.1997.9.8.1735.

[8] S. Schmid, H. Winkler, Hybrid production management system in the context of industry 4.0, in: 2022 IEEE International Conference on Industrial Engineering and Engineering Management, IEEM, IEEE, 2022, pp. 1573–1577, http://dx.doi.org/10.1109/IEEM55944.2022.9990000.

[9] M.M. Rathore, S.A. Shah, D. Shukla, E. Bentafat, S. Bakiras, The role of ai, machine learning, and big data in digital twinning: A systematic literature review, challenges, and opportunities, IEEE Access 9 (2021) 32030–32052, http://dx.doi.org/10.1109/ACCESS.2021.3060863.

[10] K. Bartsch, A. Pettke, A. Hübert, J. Lakämper, F. Lange, On the digital twin application and the role of artificial intelligence in additive manufacturing: A systematic review, J. Phys.: Mater. 4 (3) (2021) 032005, http://dx.doi.org/10.1088/2515-7639/abf3cf.

[11] W. Kritzinger, M. Karner, G. Traar, J. Henjes, W. Sihn, Digital twin in manufacturing: A categorical literature review and classification, Ifac-PapersOnline 51 (11) (2018) 1016–1022, http://dx.doi.org/10.1016/j.ifacol.2018.08.474.

[12] C. Semeraro, M. Lezoche, H. Panetto, M. Dassisti, Digital twin paradigm: A systematic literature review, Comput. Ind. 130 (2021) 103469, http://dx.doi.org/10.1016/j.compind.2021.103469.

[13] K.Y.H. Lim, P. Zheng, C.-H. Chen, A state-of-the-art survey of digital twin: techniques, engineering product lifecycle management and business innovation perspectives, J. Intell. Manuf. 31 (2020) 1313–1337, http://dx.doi.org/10.1007/s10845-019-01512-w.

[14] M. Grieves, Digital twin: manufacturing excellence through virtual factory replication, White Pap. 1 (2014) (2014) 1–7, URL https://www.3ds.com/fileadmin/PRODUCTS-SERVICES/DELMIA/PDF/Whitepaper/DELMIA-APRISO-Digital-Twin-Whitepaper.pdf.

[15] M. Grieves, J. Vickers, Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems, in: Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches, Springer, 2017, pp. 85–113.

[16] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444, http://dx.doi.org/10.1038/nature14539.

[17] L.P. Kaelbling, M.L. Littman, A.W. Moore, Reinforcement learning: A survey, J. Artif. Intell. Res. 4 (1996) 237–285, http://dx.doi.org/10.1613/jair.301.

[18] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature 323 (6088) (1986) 533–536, http://dx.doi.org/10.1038/323533a0.

[19] Y. LeCun, Y. Bengio, et al., Convolutional networks for images, speech, and time series, Handb. Brain Theory Neural Netw. 3361 (10) (1995) 1995, URL https://citeseerx.ist.psu.edu/document?doi=e26cc4a1c717653f323715d751c8dea7461aa105.

[20] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, http://dx.doi.org/10.1023/A:1010933404324.

[21] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (1995) 273–297, http://dx.doi.org/10.1007/BF00994018.

[22] V. Mnih, K. Kavukcuoglu, D. Silver, A.A. Rusu, J. Veness, M.G. Bellemare, A. Graves, M. Riedmiller, A.K. Fidjeland, G. Ostrovski, et al., Human-level control through deep reinforcement learning, Nature 518 (7540) (2015) 529–533, http://dx.doi.org/10.1038/nature14236.

[23] B. Kitchenham, S. Charters, et al., Guidelines for performing systematic literature reviews in software engineering, 2007, URL https://www.researchgate.net/publication/302924724_Guidelines_for_performing_Systematic_Literature_Reviews_in_Software_Engineering.

[24] J. Um, J. Popper, M. Ruskowski, Modular augmented reality platform for smart operator in production environment, in: 2018 IEEE Industrial Cyber-Physical Systems, ICPS, IEEE, 2018, pp. 720–725, http://dx.doi.org/10.1109/ICPHYS.2018.8390796.

[25] Y. Tai, L. Zhang, Q. Li, C. Zhu, V. Chang, J.J. Rodrigues, M. Guizani, Digital-twin-enabled IoMT system for surgical simulation using rAC-GAN, IEEE Internet Things J. 9 (21) (2022) 20918–20931, http://dx.doi.org/10.1109/JIOT.2022.3176300.

[26] D. Gupta, O. Kayode, S. Bhatt, M. Gupta, A.S. Tosun, Hierarchical federated learning based anomaly detection using digital twins for smart healthcare, in: 2021 IEEE 7th International Conference on Collaboration and Internet Computing, CIC, IEEE, 2021, pp. 16–25, http://dx.doi.org/10.1109/CIC52973.2021.00013.

[27] H. Latif, G. Shao, B. Starly, A case study of digital twin for a manufacturing process involving human interactions, in: 2020 Winter Simulation Conference, WSC, IEEE, 2020, pp. 2659–2670, http://dx.doi.org/10.1109/WSC48552.2020.9383897.

[28] Y. Shi, W. Shen, L. Wang, F. Longo, L. Nicoletti, A. Padovano, A cognitive digital twins framework for human-robot collaboration, Procedia Comput. Sci. 200 (2022) 1867–1874, http://dx.doi.org/10.1016/j.procs.2022.01.387.

[29] A. Gallala, A.A. Kumar, B. Hichri, P. Plapper, Digital twin for human–robot interactions by means of industry 4.0 enabling technologies, Sensors 22 (13) (2022) 4950, http://dx.doi.org/10.3390/s22134950.

[30] F. Pires, B. Ahmad, A.P. Moreira, P. Leitão, Recommendation system using reinforcement learning for what-if simulation in digital twin, in: 2021 IEEE 19th International Conference on Industrial Informatics, INDIN, IEEE, 2021, pp. 1–6, http://dx.doi.org/10.1109/INDIN45523.2021.9557372.

[31] B.R. Barricelli, E. Casiraghi, J. Gliozzo, A. Petrini, S. Valtolina, Human digital twin for fitness management, IEEE Access 8 (2020) 26637–26664, http://dx.doi.org/10.1109/ACCESS.2020.2971576.

[32] Z. Zhang, Y. Zeng, H. Liu, C. Zhao, F. Wang, Y. Chen, Smart DC: an AI and digital twin-based energy-saving solution for data centers, in: NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium, IEEE, 2022, pp. 1–6, http://dx.doi.org/10.1109/NOMS54207.2022.9789853.

[33] M. Matulis, C. Harvey, A robot arm digital twin utilising reinforcement learning, Comput. Graph. 95 (2021) 106–114, http://dx.doi.org/10.1016/j.cag.2021.01.011.

[34] H. Zhou, C. Yang, Y. Sun, Intelligent ironmaking optimization service on a cloud computing platform by digital twin, Engineering 7 (9) (2021) 1274–1281, http://dx.doi.org/10.1016/j.eng.2021.04.022.

[35] N. Kharlamova, C. Træholt, S. Hashemi, A digital twin of battery energy storage systems providing frequency regulation, in: 2022 IEEE International Systems Conference, SysCon, IEEE, 2022, pp. 1–7, http://dx.doi.org/10.1109/SysCon53536.2022.9773919.

[36] H.-H. Benzon, X. Chen, L. Belcher, O. Castro, K. Branner, J. Smit, An operational image-based digital twin for large-scale structures, Appl. Sci. 12 (7) (2022) 3216, http://dx.doi.org/10.3390/app12073216.

[37] Q. Song, S. Lei, W. Sun, Y. Zhang, Adaptive federated learning for digital twin driven industrial internet of things, in: 2021 IEEE Wireless Communications and Networking Conference, WCNC, IEEE, 2021, pp. 1–6, http://dx.doi.org/10.1109/WCNC49053.2021.9417370.

[38] I. Al-Zyoud, F. Laamarti, X. Ma, D. Tobón, A. El Saddik, Towards a machine learning-based digital twin for non-invasive human bio-signal fusion, Sensors 22 (24) (2022) 9747, http://dx.doi.org/10.3390/s22249747.

[39] K.S.S. Alamin, Y. Chen, E. Macii, M. Poncino, S. Vinco, A machine learning-based digital twin for electric vehicle battery modeling, in: 2022 IEEE International Conference on Omni-Layer Intelligent Systems, COINS, IEEE, 2022, pp. 1–6, http://dx.doi.org/10.1109/COINS54846.2022.9854960.
[40] F. Boulfani, X. Gendre, A. Ruiz-Gazen, M. Salvignol, Anomaly detection for aircraft electrical generator using machine learning in a functional data framework, in: 2020 Global Congress on Electrical Engineering, GC-ElecEng, IEEE, 2020, pp. 27–32, http://dx.doi.org/10.23919/GC-ElecEng48342.2020.9285984.
[41] M. Li, J. Gao, C. Zhou, X. Shen, W. Zhuang, Digital twin-driven computing resource management for vehicular networks, in: GLOBECOM 2022-2022 IEEE Global Communications Conference, IEEE, 2022, pp. 5735–5740, http://dx.doi.org/10.1109/GLOBECOM48099.2022.10001032.
[42] J. Miao, P. Zhang, UAV visual navigation system based on digital twin, in: 2022 18th International Conference on Mobility, Sensing and Networking, MSN, IEEE, 2022, pp. 865–870, http://dx.doi.org/10.1109/MSN57253.2022.00140.
[43] B. Li, Y. Liu, L. Tan, H. Pan, Y. Zhang, Digital twin assisted task offloading for aerial edge computing and networks, IEEE Trans. Veh. Technol. 71 (10) (2022) 10863–10877, http://dx.doi.org/10.1109/TVT.2022.3182647.
[44] Z. Wang, R. Gupta, K. Han, H. Wang, A. Ganlath, N. Ammar, P. Tiwari, Mobility digital twin: Concept, architecture, case study, and future challenges, IEEE Internet Things J. 9 (18) (2022) 17452–17467, http://dx.doi.org/10.1109/JIOT.2022.3156028.
[45] D. Chen, N.A. AlNajem, M. Shorfuzzaman, Digital twins to fight against COVID-19 pandemic, Internet Things Cyber-Phys. Syst. 2 (2022) 70–81.
[46] S.R. Chhetri, S. Faezi, A. Canedo, M.A.A. Faruque, QUILT: Quality inference from living digital twins in IoT-enabled manufacturing systems, in: Proceedings of the International Conference on Internet of Things Design and Implementation, 2019, pp. 237–248, http://dx.doi.org/10.1145/3302505.3310085.
[47] A. Mozo, A. Karamchandani, S. Gómez-Canaval, M. Sanz, J.I. Moreno, A. Pastor, B5GEMINI: AI-driven network digital twin, Sensors 22 (11) (2022) 4106, http://dx.doi.org/10.3390/s22114106.
[48] D. Zhang, X. Gao, A digital twin dosing system for iron reverse flotation, J. Manuf. Syst. 63 (2022) 238–249, http://dx.doi.org/10.1016/j.jmsy.2022.03.006.
[49] A. Wooley, D.F. Silva, J. Bitencourt, When is a simulation a digital twin? A systematic literature review, Manuf. Lett. 35 (2023) 940–951, http://dx.doi.org/10.1016/j.mfglet.2023.08.014.
[50] X. Wu, L. Xiao, Y. Sun, J. Zhang, T. Ma, L. He, A survey of human-in-the-loop for machine learning, Future Gener. Comput. Syst. 135 (2022) 364–381, http://dx.doi.org/10.1016/j.future.2022.05.014.
[51] X. Xu, Y. Lu, B. Vogel-Heuser, L. Wang, Industry 4.0 and industry 5.0—Inception, conception and perception, J. Manuf. Syst. 61 (2021) 530–535.
[52] F. Tao, Q. Qi, L. Wang, A. Nee, Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: Correlation and comparison, Engineering 5 (4) (2019) 653–661, http://dx.doi.org/10.1016/j.eng.2019.01.014.
[53] K.E. Harper, S. Malakuti, C. Ganz, Digital twin architecture and standards, 2019, http://dx.doi.org/10.25607/OBP-1860.
[54] Y. Zhao, N. Wang, Z. Liu, An established theory of digital twin model for tunnel construction safety assessment, Appl. Sci. 12 (23) (2022) 12256, http://dx.doi.org/10.3390/app122312256.
[55] T. Ritto, F. Rochinha, Digital twin, physics-based model, and machine learning applied to damage detection in structures, Mech. Syst. Signal Process. 155 (2021) 107614, http://dx.doi.org/10.1016/j.ymssp.2021.107614.
[56] D.-G.J. Opoku, S. Perera, R. Osei-Kyei, M. Rashidi, Digital twin application in the construction industry: A literature review, J. Build. Eng. 40 (2021) 102726, http://dx.doi.org/10.1016/j.jobe.2021.102726.
[57] Y. Zhang, K. Song, Y. Sun, S. Tan, M. Udell, "Why should you trust my explanation?" understanding uncertainty in LIME explanations, 2019, arXiv preprint arXiv:1904.12991.

Tim Kreuzer is a Doctoral Student in the program of Computer and Systems Sciences at the Department of Computer and Systems Sciences of Stockholm University. His research focuses on Machine Learning methods and Digital Twins, specializing in the integration of the two fields. He holds a bachelor's degree in Mobile Computing from Hof University (2020) and a master's degree in Software Engineering from the University of Bolzano (2022). Within the field of machine learning, his interests focus on forecasting and classifying temporal data as well as model explainability. Goal of Tim's research is the combination of AI methods with DT models in domains such as building information modeling.

Panagiotis Papapetrou is a Professor of Computer Science at the Department of Computer and Systems Sciences at Stockholm University. His area of expertise is algorithmic data mining with a particular focus on mining and indexing sequential data, complex metric and non-metric spaces, biological sequences, time series, and sequences of temporal intervals. Panagiotis received his Ph.D. in Computer Science at Boston University in 2009, was a postdoctoral researcher at Aalto University during 2009-2013, and lecturer at the University of London during 2012-2013. He has participated in several EU Projects, NSF grants, and Academy of Finland centers of excellence. He also holds a starting grant (Etableringsbidrag) from the Swedish Research Council. He has served as general chair of the 15th International Symposium on Intelligent Data Analysis (IDA) 2016 and 2024, has participated in the co-organization of three Workshops in the area of Data Mining and Knowledge Discovery, and board member of the Swedish AI Society. He is Associate Editor for the Journal of Data Mining and Knowledge Discovery.

Jelena Zdravkovic is a Professor at the Department of Computer and Systems Sciences (DSV) at Stockholm University. She has a Ph.D. in Computer and Systems Sciences at The Royal Institute of Technology (KTH) in 2006. Jelena has published over 100 refereed papers on the topic of IS engineering in international conferences and scientific journals. Her current research activities include data-driven requirements engineering with a focus on digital ecosystems and digital twins. Jelena has participated in several national and international projects related to requirements for IS, and model-driven IS development. Jelena was on the editorial board of the BISE journal for about 10 years, and is currently associate editor for the RE Journal, and guest editor for the SoSYM journal; as well as serving as a regular reviewer for them, and others, including Elsevier's Journal of Systems and Software, the Information and Software Technology Journal, and IEEE Computing. She has served as the program chair and is on the program board of several international conferences, including CAiSE, ER, RCIS, PoEM, and as a program committee member in a number of other conferences and workshops in the IS engineering discipline.