


Review

Navigating the Evolution of Digital Twins Research through Keyword Co-Occurrence Network Analysis

Wei Li, Haozhou Zhou, Zhenyuan Lu and Sagar Kamarthi * 

Department of Mechanical and Industrial Engineering, Northeastern University, Boston, MA 02115, USA; li.wei10@northeastern.edu (W.L.); zhou.haoz@northeastern.edu (H.Z.); lu.zhenyua@northeastern.edu (Z.L.)

* Correspondence: s.kamarthi@northeastern.edu

Abstract: Digital twin technology has become increasingly popular and has revolutionized data integration and system modeling across various industries, such as manufacturing, energy, and healthcare. This study aims to explore the evolving research landscape of digital twins using Keyword Co-occurrence Network (KCN) analysis. We analyze metadata from 9639 peer-reviewed articles published between 2000 and 2023. The results unfold in two parts. The first part examines trends and keyword interconnection over time, and the second part maps sensing technology keywords to six application areas. This study reveals that research on digital twins is rapidly diversifying, with focused themes such as predictive and decision-making functions. Additionally, there is an emphasis on real-time data and point cloud technologies. The advent of federated learning and edge computing also highlights a shift toward distributed computation, prioritizing data privacy. This study confirms that digital twins have evolved into complex systems that can conduct predictive operations through advanced sensing technologies. The discussion also identifies challenges in sensor selection and empirical knowledge integration.

Keywords: digital twins (DT); keyword co-occurrence network (KCN); artificial intelligence (AI); sensors; scientometric analysis



Citation: Li, W.; Zhou, H.; Lu, Z.; Kamarthi, S. Navigating the Evolution of Digital Twins Research through Keyword Co-Occurrence Network Analysis. *Sensors* **2024**, *24*, 1202. <https://doi.org/10.3390/s24041202>

Academic Editor: Gianluigi Ferrari

Received: 7 January 2024

Revised: 3 February 2024

Accepted: 7 February 2024

Published: 12 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The concept of digital twins has evolved beyond its original role in product lifecycle management [1] and become an essential element in the digital transformation across various sectors. The digital twin applications typically involve the creation, utilization, and sustainment of a virtual counterpart of a physical system, facilitating real-time, two-way data exchanges [2]. Digital twins enhance human-machine interactions and inter-machine communications. They dynamically and behaviorally mirror their physical counterparts, integrating both raw and processed data to reflect real-world conditions accurately. While proactive development of digital twins is advocated for optimal integration, retrofitting remains a common practice for existing systems [3].

Digital twins are categorized into four functional levels: representation, replication, reality, and relational [4], as shown in Figure 1. The foundational level, *representation*, focuses on data collection and physical system representation. The form of digital twins at this level is usually real-time data connectivity and visualization. A virtual model is created at the *replication* level to duplicate the physical system and produce the same outputs as the physical system. With the aid of cutting-edge simulation approaches, virtual models have been capable of monitoring and controlling industrial systems with more complex configurations [5]. Digital twins at this level are usually equipped with basic analytical models that can analyze and predict system conditions given existing scenarios. The *reality* level expands the digital twin's capabilities for exploratory "what-if" analyses, enabling predictions for hypothetical changes and scenarios. The most advanced level, *relational* function, equips digital twins with machine learning models, providing insights that can be

acted upon to optimize the physical system performance. Digital twins at this level achieve a seamless close-loop bidirectional data flow and integration between the physical and virtual realms.

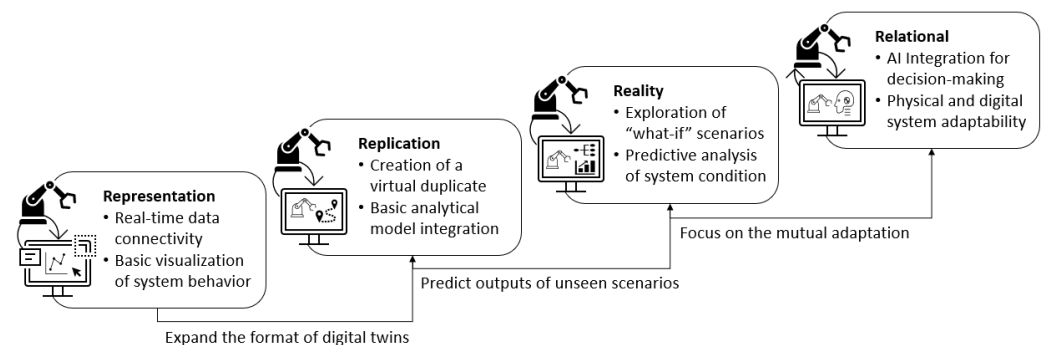


Figure 1. Four levels of functionalities of digital twins.

Smart technologies such as the Internet of Things (IoT) and Cyber-Physical Systems (CPSs) have been thoroughly studied before the emergence of digital twins. The IoT is a network of physical objects with embedded sensors and other technologies that connect and exchange data with other devices [6]. On the other hand, CPSs are complex systems that integrate the cyber world and the physical world through computing, communication, and control [7]. Although CPSs and digital twins seem similar in definition, they differ in their primary focuses. Digital twins focus on creating a comprehensive virtual model that mimics and predicts the behavior of its physical counterpart [8]. In contrast, CPSs focus on real-time control and the ability to respond to physical states, often through direct sensor and actuator involvement.

Digital twins rely on the same or similar the IoT and CPS-enabling technologies. Because of this, digital twins, as with the IoT and CPSs, have expanded beyond manufacturing and into various other industries. The sensing ecosystem, which includes traditional physical sensors, advanced data analytics, and processing platforms, is central to the expansion of digital twins [9]. The sensing ecosystem captures and interprets the vast streams of data generated by the IoT and monitored by CPSs, forming the backbone of digital twin functionality and setting the stage for the vital role of sensors in digital twin-supported integrated systems.

Sensors, commonly represented by physical devices such as accelerometers and temperature gauges, are the fundamental interface between the physical and digital worlds. The basic structure of a sensor module serves to convert measurable physical phenomena into data streams for analysis and application in digital systems [10]. Yet, the definition of sensors has expanded in today's interconnected environment. Today, a sensor can be anything that translates real-world variables into data, ranging from social media posts that gauge public sentiment to medical tests that provide insights into a patient's health. This broader interpretation of sensors enables sensors to serve as the primary medium for data flow and analysis across various applications.

Table 1 features a collection of papers that review digital twin applications in various areas highlighting their contributions and identifying gaps and opportunities. This work adds quantitative research trend analysis to the current digital twin review landscape with Keyword Co-occurrence Network (KCN) analysis. KCN analysis is a tool to analyze the research landscape from the metadata of literature [11]. This method allows us to thoroughly examine and interpret the extensive range of digital twin research.

Table 1. A selection of digital twins review papers.

Application Area	Key Contributes	Gaps/Opportunities Identified	Ref.
General Applications	Reviews covering concepts, key enabling technologies, and implementation of digital twins, including challenges and prospects across multiple domains	The need for standardization, data availability, processing power, interdisciplinary collaboration, and development of reference frameworks and performance metrics	[12–15]
Smart Manufacturing	Reviews digital twins integration in Industry 4.0 and digital supply chains and their optimization potential	The need for design frameworks, early detection of design flaws, clarity in research focus, and organized research environment	[16,17]
Smart Grid and Smart City	Review of digital twins in energy management and infrastructure durability	Challenges in data management, analysis, real-time interaction, and effective distributed sensing updating	[18,19]
Agriculture	Current trends, roadmap, and open questions in digital twins for agriculture	The need for automated decision-making support and complex examples	[20–22]
Smart Healthcare	Review of digital twin applications in precision medicine, clinical trial design, hospital operations, and platforms supporting mobile health applications	Technical, regulatory, and ethical challenges in healthcare digital twins	[23,24]
Education	Review of digital twins in remote and virtual laboratories	Integration of digital twin concepts into educational systems	[25]

This review paper serves digital twin researchers and architects. The KCN method reveals the interconnectedness of knowledge components, concepts, technologies, and methodologies in digital twin research, aiding researchers in identifying emerging trends and under-researched areas. For digital twin architects, the KCN analysis provides information on the practical application of sensors and other advanced digital twin technologies. It can assist them in making informed architectural decisions and understanding the evolving landscape of digital twin applications in relation to Cyber-Physical Systems.

The remainder of this work is structured as follows. In the methodology section, we explain the process of KCN analysis and its implementation in the context of the digital twin literature. In the results section, we present the analysis results, which include the temporal analysis of research trends, mapping of sensing technology to application fields, and detailed analysis of digital twin applications in various fields. These insights are presented with a series of visualizations and tables that illustrate the interconnected landscape of digital twin research. In the discussion section, we explain the implications of these findings and conclude by reflecting on the challenges currently being faced and the potential paths for the future development of digital twins.

2. Methods

This study applies KCN analysis to investigate research trends in digital twin technology. The methodology includes a temporal KCN analysis to identify trends over time and a detailed review of principal application categories. This section outlines the process for article collection, keyword extraction, KCN construction, and network evaluation metrics.

2.1. Article Collection and Screening

This study began with a thorough search for literature related to recent advancements in digital twins. We queried literature from Engineering Village, IEEE Xplore, and PubMed. Engineering Village and IEEE Xplore ensure comprehensive coverage in engineering-related subjects, and PubMed provides additional coverage of medical literature. We selected articles that contain the terms “digital twin” OR “digital twins” in the metadata (title, keywords, and abstract). After narrowing the search to peer-reviewed journal articles and conference proceedings published in English from 2000 to 2023, we identified a total of 9639 papers and downloaded the metadata of this article collection.

We classified the papers according to their respective application categories to analyze the research trends of digital twins and their applications in different fields. We identified six primary application categories. Figure 2 displays the subtopics under each primary category. Although a paper may fall under multiple categories, it is assigned to the most relevant category based on its content.

Fundamental Research <ul style="list-style-type: none"> - Sensor advancement - Data science - Basic theories - Application-agonistic topics 	Manufacturing and Supply Chain <ul style="list-style-type: none"> - Manufacturing - Cyber-physical system - Control system and robotics - Production and life-cycle 	Energy and Power Grid <ul style="list-style-type: none"> - Smart grid - Power system - Energy management - Battery technology
Healthcare and Human-centric <ul style="list-style-type: none"> - Healthcare - Education - Human services - Recreation 	Infrastructure and Transportation <ul style="list-style-type: none"> - Smart city - Smart farming - Transportation sector (e.g., automotive and aerospace) 	Business and Asset Management <ul style="list-style-type: none"> - Business management - Asset management - Social issue alleviation - Defense applications

Figure 2. Six primary digital twin application categories.

2.2. Keyword Co-Occurrence Network Construction

After collecting the metadata from the publications identified in the initial screening, we converted this unstructured data into a structured format suitable for quantitative temporal analysis. To achieve this, we started by extracting keywords and key phrases from the abstract, keywords section, and the title of each paper. We then used a Natural Language Processing (NLP) toolkit to extract essential information while minimizing language biases [26]. The toolkit automatically broke down the title and keyword strings into phrases, eliminated common words such as “a” and “the”, reduced words to their basic form, and reconciled different terminologies referring to the same concept, such as “cyber-physical systems” and “CPS”.

To construct a KCN using the structured data, each keyword is considered as a node, and the co-occurrence of a pair of keywords in the same paper is treated as an edge connecting the co-occurring keyword pair (node pair). The resulting KCN is undirected and weighted, with edge weights indicating co-occurrence frequencies. The KCN, which consists of n unique keywords, is stored in an $n \times n$ adjacency matrix \mathbf{a} . The value of each cell in the matrix, a_{ij} , is set to 1 if a connection exists between keyword i and j , and 0 otherwise.

This study conducts two types of analyses using the KCN. The first type is a temporal research trend analysis, in which we segment publications into distinct time windows: 2000–2020, 2021, 2022, and 2023. We build a separate KCN for each time window to capture the evolving trends over time. The second analysis focuses on capturing research highlights within each application field. Therefore, we divided the publications by their application category and constructed an individual KCN for each category. After constructing the network, we calculated various network metrics for our subsequent analyses.

2.3. Network Metrics

This study applies five network metrics to evaluate the KCN. These metrics help identify important keywords, understand their interconnections, and determine the overall structure and trend of the research field. These metrics measure node centrality, connections, and the local topology of node groups.

Node centrality is measured by degree and strength. The **degree** of a node refers to the number of unique nodes it directly connects to. The degree can be calculated using the adjacency matrix \mathbf{a} . As shown in Equation (1), the degree of node i is the sum of a_{ij} . Node j belongs to the group of nodes N_i that directly connects to node i .

$$d_i = \sum_{j \in N_i} a_{ij} \quad (1)$$

The **strength** of a node counts the number of connections it has, taking into account the frequency of co-occurrences. As shown in Equation (2), the strength of node i is a weighted sum of a_{ij} , where w_{ij} is the number of connections between node i and node j .

$$s_i = \sum_{j \in N_i} a_{ij} w_{ij} \quad (2)$$

Average weight as a function of endpoint degree quantifies the relationship between the connectivity of nodes and the strength of their co-occurrences. To calculate this, we define the endpoint degree of an edge connecting node i and j as $d_i d_j$. We then examine the relationship between w_{ij} and $d_i d_j$ for every edge in the network. Due to the multiplicity of edges with identical endpoint degree values in large networks (e.g., 1×50 and 5×10 both equal 50 and hence will have the same endpoint degree), we aggregate edges into set E when they share the same endpoint degree. The average weight for edges in set E is then calculated by Equation (3), where $|E|$ is the count of edges in set E . By plotting \bar{w}_E against the endpoint degree of edge set E , we can visualize the patterns in keyword connectivity. A positive correlation indicates that highly connected keywords (i.e., keywords connected with high strength) tend to co-occur more, while a negative correlation suggests that less connected keywords (i.e., keywords connected with low strength) tend to co-occur more.

$$\bar{w}_E = \frac{\sum_{(i,j) \in E} w_{ij}}{|E|} \quad (3)$$

The local topology of a network was measured by the average weighted nearest neighbor degree and weighted clustering coefficient. The **average weighted nearest neighbor degree** indicates the strength of a node's connection with its high- or low-degree neighbors. As shown in Equation (4), the average weighted nearest neighbor degree d_i^w of a node i is obtained by adding up the weighted connections of a node's direct neighbors and dividing it by the node's strength. A higher value indicates that a keyword is typically associated with highly connected keywords. By plotting d_i^w against d_i , we can visualize how nodes of different degrees behave in terms of connectivity.

$$d_i^w = \frac{1}{s_i} \sum_{j \in N_i} a_{ij} w_{ij} d_j \quad (4)$$

The **weighted clustering coefficient** measures the level of interconnectedness among a node's neighboring nodes, taking into account the weight of each connection. As shown in Equation (5), the coefficient C_i^w is calculated by averaging the weights w_{ij} and w_{ih} of the connections that node i shares with its neighbors, node j and node h . The calculation incorporates a normalization factor $s_i(d_i - 1)$, which adjusts for the number of potential connections and the strength of each. A high C_i^w value indicates that a node's neighbors are not only interconnected but also connected through stronger ties; this suggests a more cohesive and tightly knit structure around the node. In the context of KCN, a high-weighted clustering coefficient for a keyword indicates robust thematic clustering.

$$C_i^w = \frac{1}{s_i(d_i - 1)} \sum_{j,h \in N_i} \left(\frac{w_{ij} + w_{ih}}{2} \right) a_{ij} a_{ih} a_{jh} \quad (5)$$

3. Results

In this section, we first analyze the evolution of digital twins research over time, featuring visualizations of emerging and declining topics. Then, we examine the application of sensor technology across different fields, highlighting principal keywords in each area. Finally, we present specific case studies from each application field, showcasing the practical implementations of digital twin technology.

3.1. Research Landscape Evolution Over Time

From the KCN analysis results, we see a clear growth and diversification trend in the digital twin field. Table 2 presents the statistics of KCNs from 2000 to 2023. Here, we observe a substantial rise in the number of articles, keywords, and links, particularly after 2020. The increase in articles indicates a surge in research activities, while the growth in links points to an expanding web of interconnected topics.

Table 2. Temporal analysis of four KCNs from four time periods.

Metric	2000–2020	2021	2022	2023
Number of Articles	1462	1561	3041	3575
Number of Keywords	1060	1148	2346	2770
Number of Links	6082	7382	19,634	26,299
Average Network Strength	14.85	16.47	21.22	24.53
Max Strength	1665	2092	4658	7284
Average Network Degree	11.48	12.86	16.74	18.99
Max Degree	652	807	1709	2305
Average Network Weight	1.29	1.28	1.27	1.29
Max Weight	77	72	139	158

The development stage of this research field can also be assessed with the K value, based on Kuhn's model of scientific progression [27]. The K value is calculated by dividing the number of unique keywords by the frequency of those keywords within a discipline. Derived from Table 2, the K values for four different time windows are 0.174, 0.155, 0.119, and 0.105, respectively. The declining trend in the K value, inversely proportional to the growing number of publications, suggests that the field of digital twins is in the midst of an evolution, aligning with Kuhn's pre-revolution or revolution stage.

Figure 3 reinforces this observation by showing the distribution of articles, keywords, and links across the four time periods, with significant growth in the latter two years. This suggests not only an increase in the research volume but also expansion in the complexity within the field. Figure 4 expands on these data by comparing the average network strength and the maximum weight of the network, both of which indicate an increase in inter-article and inter-topic connections.

Figure 5 offers a distribution of keyword degrees, strengths, and link weights. The upward trends in average and maximum network degrees from Figures 4 and 5 hint at a broadening scope of individual topics and articles, suggesting an increasingly collaborative research environment where topics are more interconnected. Notably, the outliers represent the keywords that are highly connected and centric to this research field. We will present and discuss these topics in the following sections.

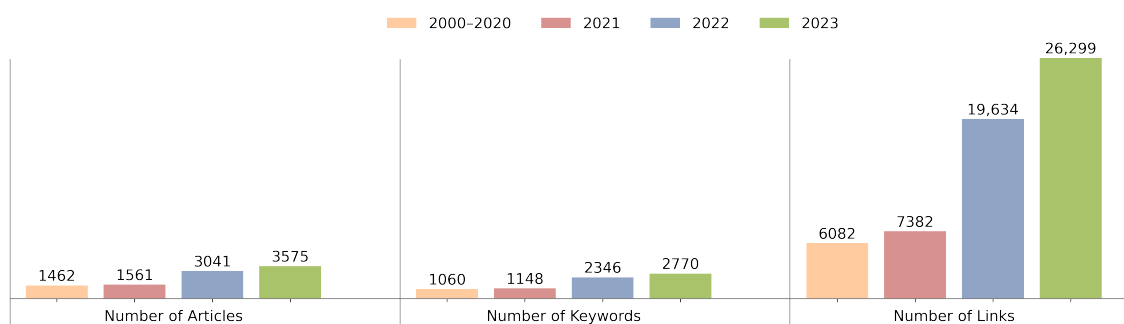


Figure 3. Number of articles, keywords, and links over the four time periods.

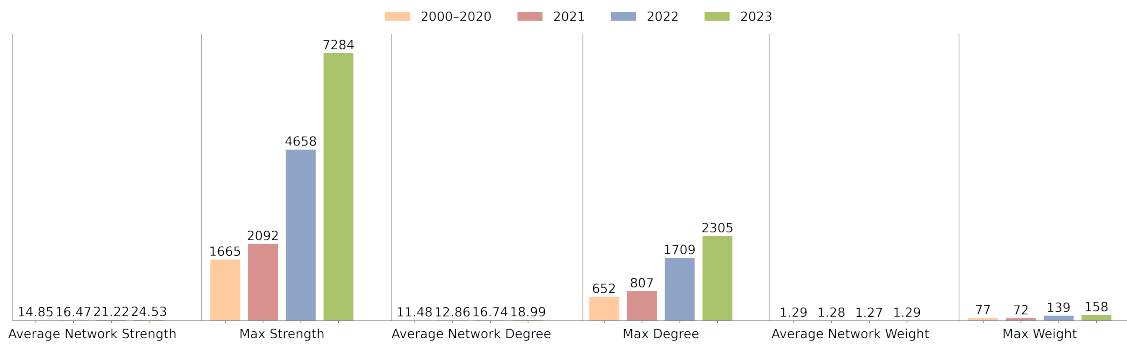


Figure 4. Growth trends in KCN parameters such as average network strength, maximum strength, average network degree, and maximum degree.

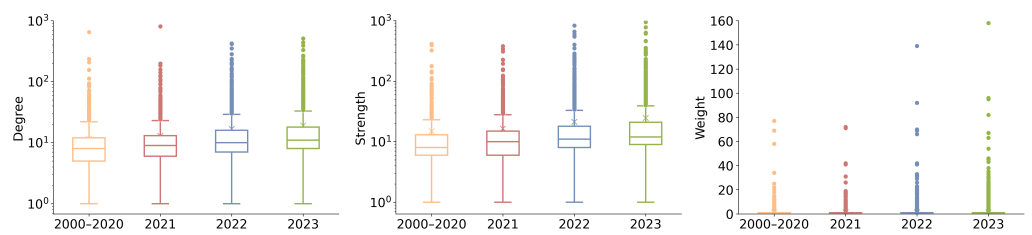


Figure 5. Boxplots of keyword degree, strength, and link weight distribution in the KCN.

Figure 6 provides various insights into the dynamics of the network. Figure 6a shows the probability density function of keyword degree. A shift toward the right over time indicates that certain keywords are becoming increasingly prominent within the network. Figure 6b examines the average weight as a function of endpoint degree. The positive linear trend suggests that keywords with a higher degree tend to form stronger connections with other keywords. However, it is not clear from this graph alone whether these connections tend to be with other highly connected keywords or emerging keywords. In addition, the subtle shift toward the right with time means that the combination of keyword degrees associated with a given average weight has been growing with time, suggesting popular keywords start to be the hubs that connect newer topics into the network, facilitating the network’s growth.

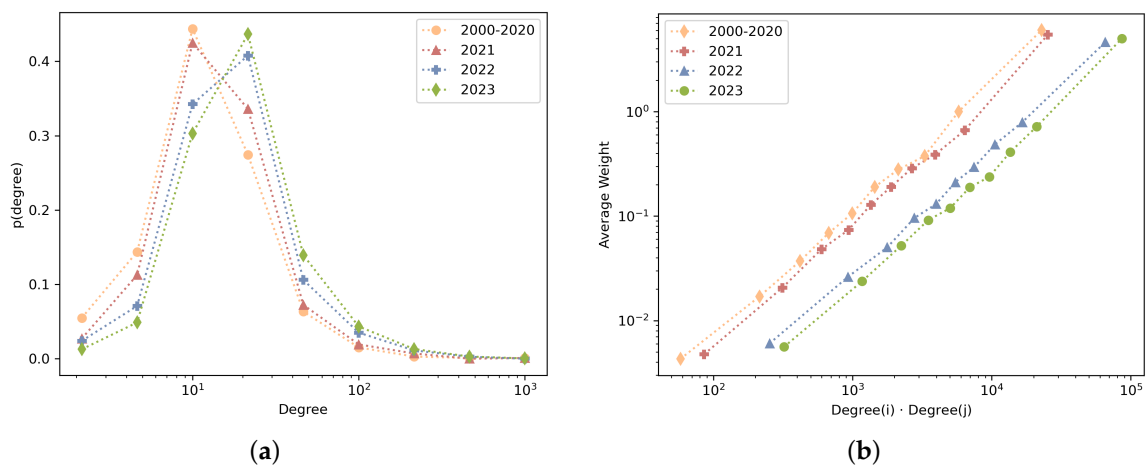


Figure 6. Cont.

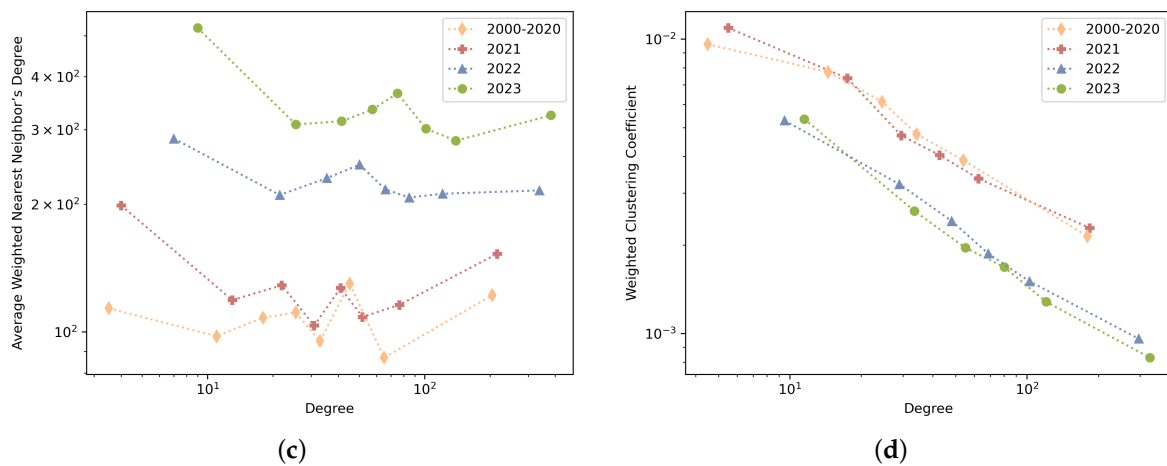


Figure 6. Linecharts of KCN metrics: (a) probability density function of keyword degree, (b) average weight as a function of endpoint degree, (c) average weight as a function of endpoint degree, and (d) weighted clustering coefficient as a function of node degree.

Figure 6c presents the relationship between the average weighted neighbor's degree and the node degree. In all four time windows, there is no clear correlation between the degree of a node and its neighbor's degree. This complements the insights from Figure 6b and shows that highly connected keywords connect with a diverse range of nodes rather than only with other highly connected nodes. To accompany this insight, Figure 6d shows the decreasing trend in the weighted clustering coefficient, indicating that highly connected nodes act more as bridges than remaining within isolated clusters, pointing to an expanding and diversifying field.

These visualizations and metrics depict the characteristics of a rapidly growing complex field, with foundational research expanding and certain topics gaining more prominence. However, the consistent average weight across years also suggests that the additional links may not always contribute to the foundational research, raising questions about the depth and influence of recent publications. This nuanced view of the field's evolution indicates both robust growth and areas requiring further investigation to understand the research impact.

3.2. Emerging and Declining Research Topics

Figures 7 and 8 trace the changes in keyword relevance over time, from the earliest time window of 2000–2020 to the most recent time window of 2023. In each time window, we ranked keywords based on their strength, which is determined by the number of connections each keyword has. We then compared the ranking of keywords from both time windows to assess emerging or declining trends.

Because there was a substantial increase in overall keyword strength from the earlier to the later time window, we used rank as a proxy for a keyword's relevance within its specific period. Additionally, we categorized the keywords into two groups: those relating to digital twin applications and those relating to the sensing ecosystem, which includes sensors, machine learning methods, and computational systems.

It is important to note that a slight decline in a keyword's rank does not necessarily indicate a decrease in its importance or research focus. Instead, it may indicate a natural transition of the keyword from a novel research area to a more established topic that no longer occupies the forefront of emerging research themes. This shift can be seen as a maturation process within the research landscape, where once-novel concepts become foundational elements of the field.

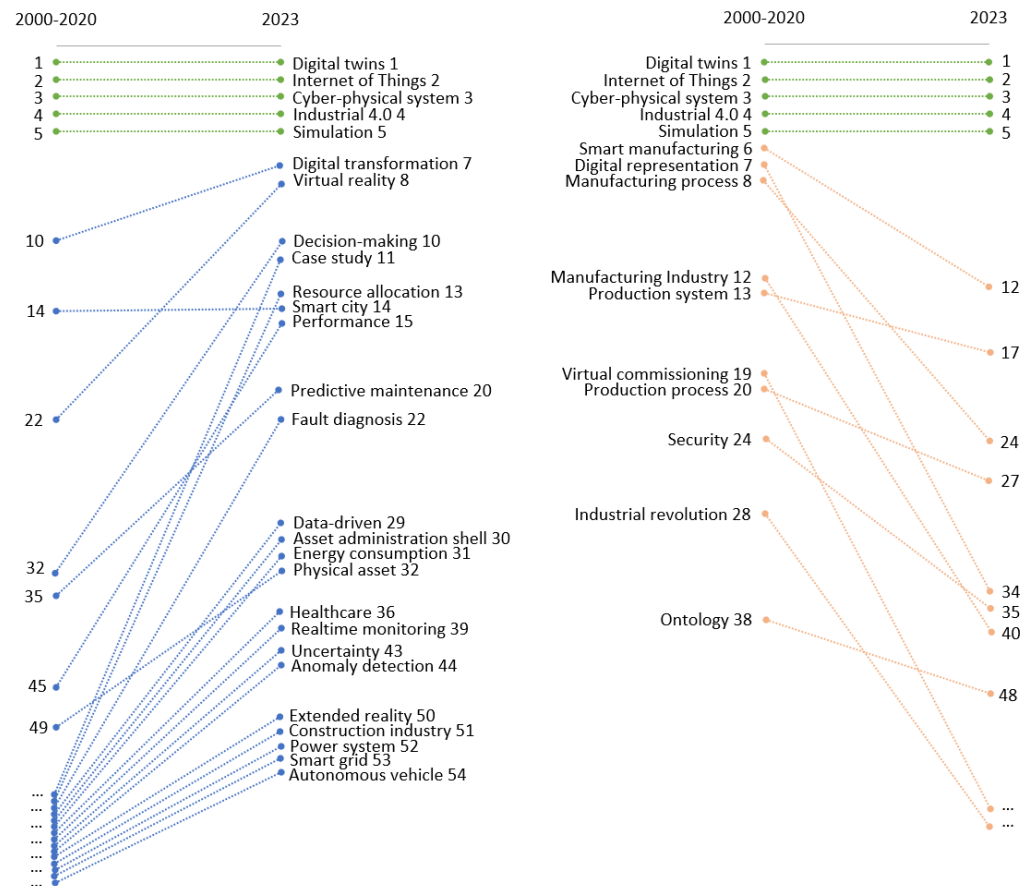


Figure 7. Emerging and declining keywords of digital twin applications from 2000–2020 to 2023.

Figure 7 presents keywords related to digital twin applications. Notably, the top five keywords, namely digital twins, Internet of Things, Cyber-Physical Systems, Industry 4.0, and simulation, have maintained the top five positions with no rank change, indicating that they have endured centrality and significance in digital twin research for over two decades. Digital twins, as the literature searching criteria, is naturally included in all research. The Internet of Things is significant for providing the sensor data that feeds digital twins. Cyber-Physical Systems are essential as they constitute the framework in which digital twins operate, integrating computation with physical processes to enable automated decision making. Industry 4.0 represents the current trend of automation and data exchange in manufacturing technologies, including Cyber-Physical Systems, the IoT, and cloud computing, which are inherently linked to the concept of digital twins. In addition, simulation serves as the analytical engine that enables the virtual representation to predict the behavior and performance of its physical counterpart.

There are two types of keywords that indicate emerging trends: application fields (areas where digital twins are being applied) and functions (what digital twins help achieve). The emerging application fields for digital twins include smart cities, energy consumption, healthcare, the construction industry, power systems, smart grids, and autonomous vehicles. The more digitalized and intelligent infrastructure in these areas enables the implementation of digital twins. The increasingly diversified application fields for digital twins also explain the slight decline of smart manufacturing and the manufacturing industry in the right panel.

The emerging functions include digital transformation, decision making, resource allocation, predictive maintenance, fault diagnosis, and real-time monitoring. The trend can be attributed to advancements in machine learning and sensor technologies. As machine learning algorithms have become more sophisticated, digital twins are now able to not only replicate physical systems but also transform and optimize them. Digital twins also have

enhanced decision-making capabilities, enabling automated and informed decisions based on predictive analytics and real-time data.

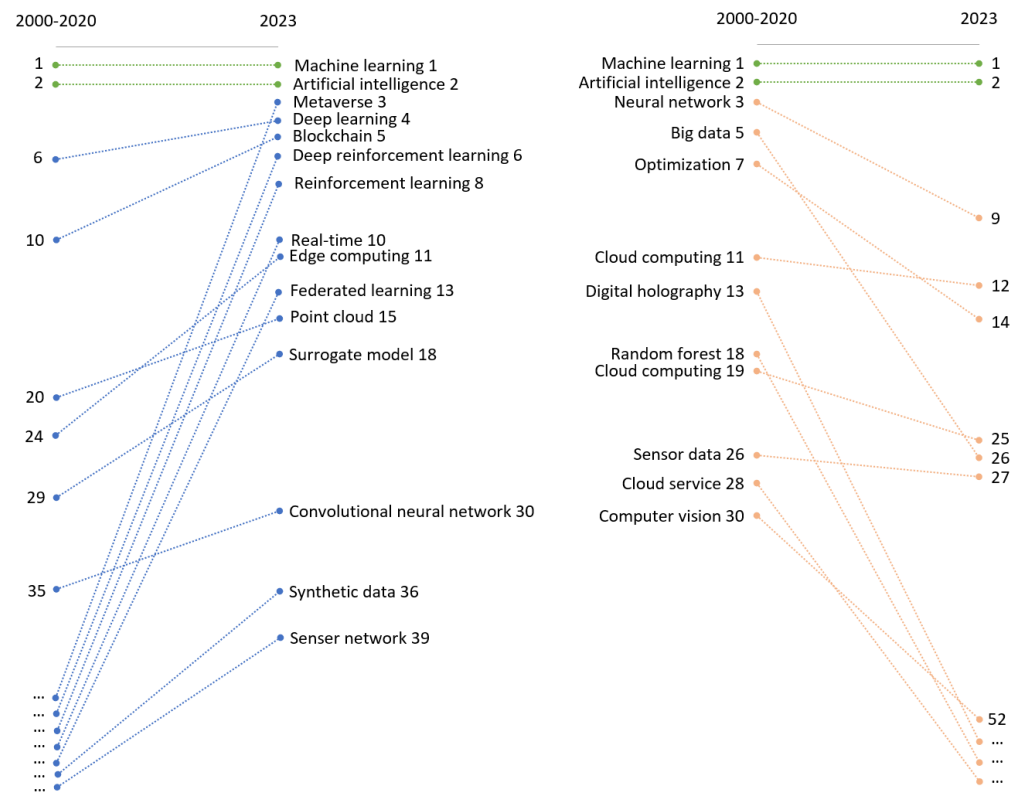


Figure 8. Emerging and declining keywords of digital twins sensing technology from 2000–2020 to 2023.

Figure 8 presents keywords related to sensor and machine learning technology. The top two keywords are machine learning and artificial intelligence. The emerging keywords related to sensors are real time, point cloud, and sensor network, highlighting the growing demand for sensors that can deliver immediate, interconnected, and diverse data types. Regarding the computation architecture that supports digital twins and machine learning functions, we notice a rising trend in edge computing and metaverse and a declining trend in cloud computing. This points to a research area pivoting toward distributed computing paradigms, suggesting a move to bring processing closer to the data source for quicker insights. This trend implies that while cloud computing has become a well-established field, the frontier of research is moving toward systems that can handle analytics at the edge of networks.

As for the machine learning-related keywords, emerging models include deep learning, reinforcement learning, federated learning, surrogate models, and convolutional neural networks. This emergence corresponds to the need for sophisticated analytical tools capable of processing complex, multimodal sensor data. These methods are particularly suited to the demands of digital twins, offering enhanced capabilities for privacy preservation and data security.

3.3. Mapping Keywords to Application Fields

The previous analysis has focused on the temporal characteristics of the research field and the trends in the relevance of the top keywords. In this section, we will delve deeper and examine how digital twin technology, specifically sensing, machine learning, and computation technologies, are being applied in different fields.

In the methodology section, we mentioned that we classified the literature into six application categories. In this section, we draw insights from each category and visualize the insights using a Sankey diagram. The left column of the Sankey diagram lists the keywords of interest, while the right column represents the application fields. The numbers on the left represent the number of papers that contain the keyword of interest. The number

on the right is a summation of all streams of numbers from the left. Since we only selected the top keywords in each category to visualize, the number on the right should not be confused with the total number of papers in each category.

Figure 9 displays the mapping from sensor technology to different digital twin application fields. Real-time data and point cloud emerge as the most prevalent keywords, which validates the trend from the slope graph. From this graph, there are two types of sensor keywords: focused sensing technologies and cross-domain technologies. As for focused keywords, process data and vibration have found their place in manufacturing settings as they are common practices for machine and equipment monitoring. Electrocardiograms and cardiac electrophysiology in healthcare may be linked to their potential to create high-fidelity visualizations for cardiac twins. LiDAR shows a strong association with infrastructure applications, likely due to its precision in capturing environmental data for smart city applications.

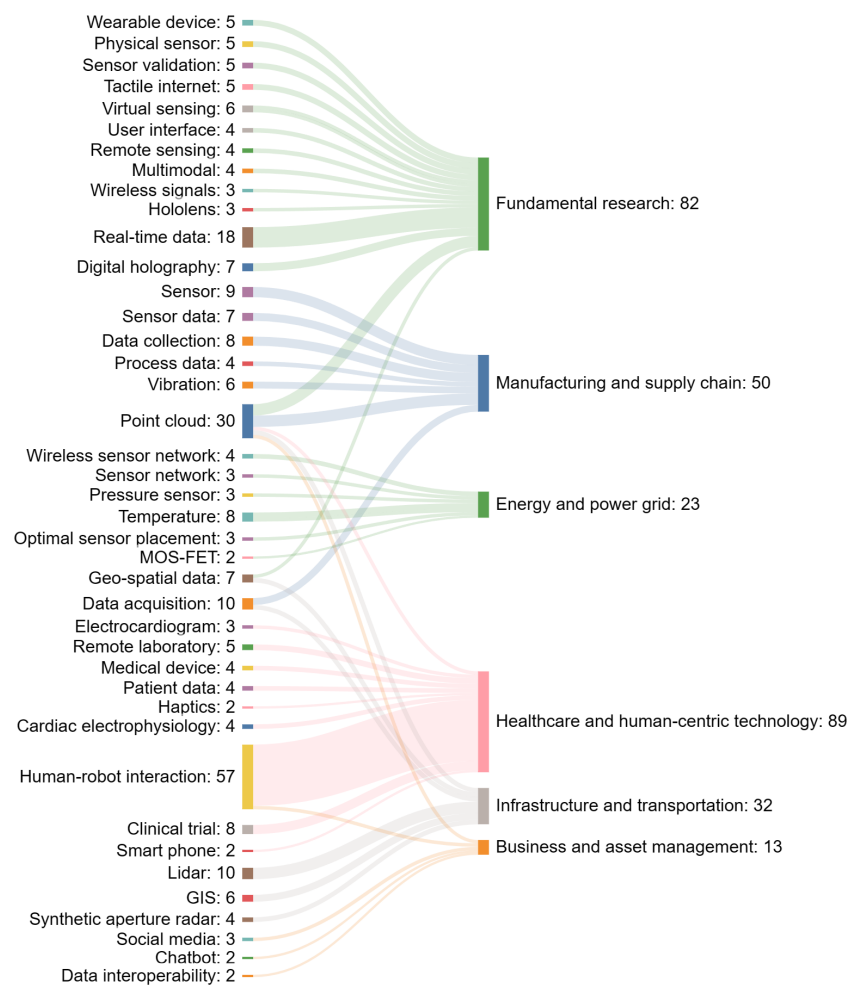


Figure 9. Mapping of sensor technology to digital twin application areas.

Cross-domain keywords such as point cloud, data acquisition, and human–robot interaction point toward the versatility of these technologies. Point cloud data, with their high-resolution spatial information, are crucial not only in manufacturing and logistics but also in infrastructure for transportation and urban planning. Data acquisition stands out as a foundational element in the sensing ecosystem to ensure the quality of high-frequency and multimodal sensing data. Human–robot interaction emphasizes the increasing collaboration between humans and automated systems. In healthcare, this could translate to robotic surgery or patient care systems, while, in manufacturing, this can pertain to collaborative robots working alongside human operators.

Figure 10 displays the mapping from machine learning methods to different digital twin application fields. Machine learning bestows active digital twins with decision-making capabilities in various applications. Neural networks and deep learning algorithms play a prominent role in pattern recognition and predictive analytics. The strong presence of reinforcement learning, particularly in fundamental research, signals an interest in developing digital twins capable of autonomous decision making and optimization—a critical feature for systems that learn and adapt over time.

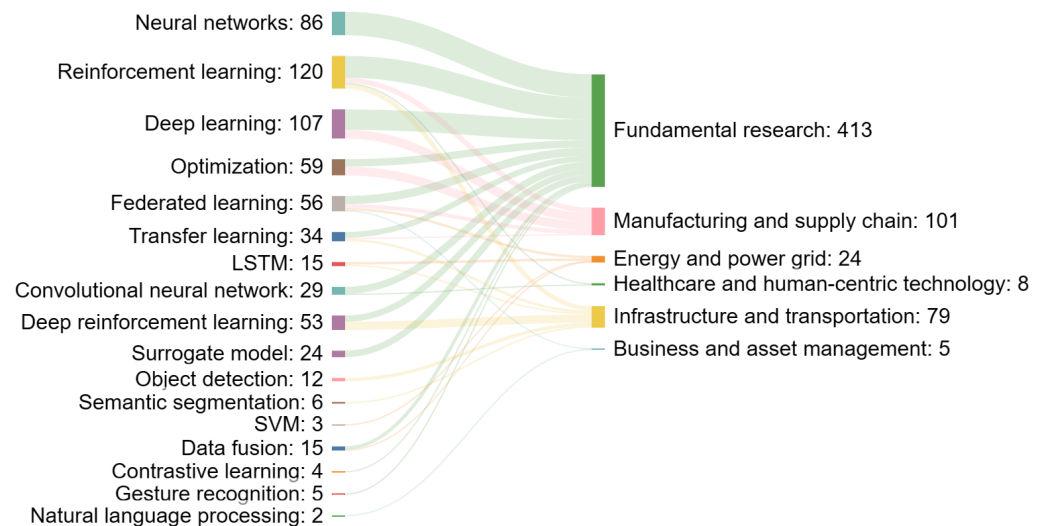


Figure 10. Mapping of machine learning methods to digital twins application fields.

The emergence of federated learning points to a growing concern for data privacy and distributed computation, enabling collaborative model training without centralized data storage. This approach aligns well with digital twins, which often require the synthesis of distributed data while sustaining confidentiality, particularly in healthcare and business settings.

The strong connection between optimization techniques and manufacturing and supply chain applications underlines the role of digital twins in process improvement and efficiency gains. Meanwhile, the intersection of convolutional neural networks with infrastructure and transportation highlights their importance in image and video processing tasks relevant to these fields.

Interestingly, the relatively modest numbers attached to healthcare and human-centric technology may reflect the nascent integration of machine learning into these regulated domains, where safety and validation are paramount.

Overall, Figure 11 displays the mapping from computational technologies to different digital twin application fields. Blockchain's notable presence across multiple fields, especially in business and asset management, highlights its role in enhancing security, transparency, and traceability. Its application within manufacturing and supply chain domains indicates its potential to revolutionize how data across the digital twin lifecycle are securely managed and shared.

The Metaverse, often associated with immersive virtual environments, shows a substantial intersection with infrastructure and transportation. This could point toward the Metaverse's capacity for sophisticated simulations and virtual testing environments, which are crucial for planning and managing large-scale infrastructural projects. As suggested by the slope chart above, cloud computing displays a slightly declining influence, indicating a shift toward distributed computing paradigms such as edge computing. Big data and data-driven keywords maintain a steady connection with fundamental research, reflecting the ongoing need to process and analyze large datasets within the digital twin sphere to extract meaningful insights. In addition, semantic interoperability and data fusion, though not as dominant, indicate niche but vital areas in ensuring that digital twins can communicate effectively across systems and synthesize information from disparate sources.

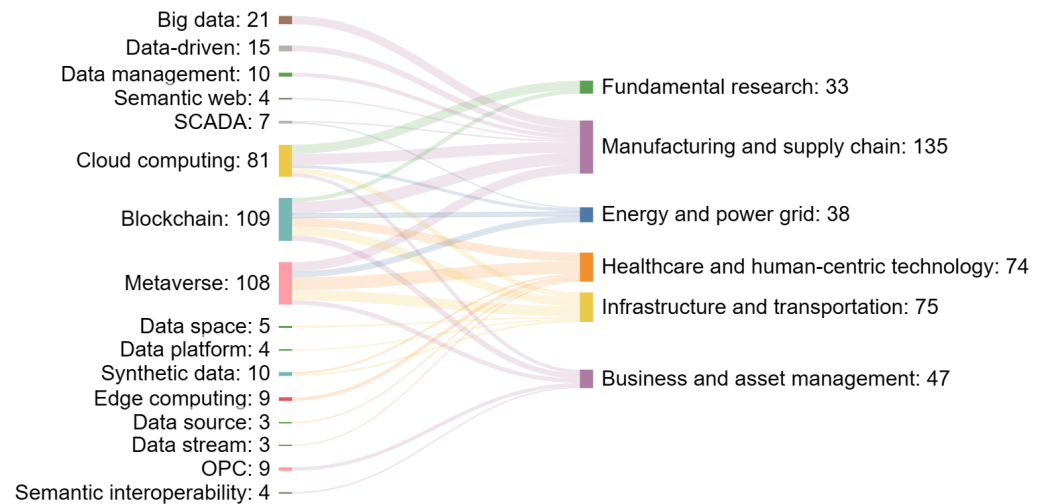


Figure 11. Mapping of computation technology to digital twin application areas.

3.4. Specific Cases in Each Application Field

Guided by the insights from the previous section, we select and review specific instances of digital twin research in this section. The tables in this section are a curated collection of publications based on the highlighted keywords from our Sankey analysis. This section aims to transition from high-level trends to individual research efforts, providing examples of how sensing, machine learning, and computation technologies are implemented within various application areas.

Table 3 presents a selection of studies in fundamental research of digital twins. The study on sensor calibration within building systems [28] demonstrates the ongoing effort to synchronize physical and virtual sensor data, which is a crucial step for accurate digital twin simulations. Research into sensor reliability [29] tackles the challenge of predictive maintenance by using redundant digital sensors to foresee potential sensor failures. Both studies emphasize the significance of sensor calibration in maintaining the operational integrity of digital twins. Challenges remain in improving the accuracy of a virtual model while maintaining the complex system built upon multiple sensors. Wearable ECG sensors [30] have been studied for low-latency signal analysis, enhancing the responsiveness of digital twins. This research resonates with the need to make digital twins interactive and user-centric. In the future, digital twins will serve not only as tools for simulation and monitoring but also as an end-to-end platform for interaction, providing intuitive feedback to users. The integration of tactile sensors in tactile devices [31] offers insight into the sensory augmentation possibility within digital twins, while the use of LiDAR for user interface design [32] highlights the importance of high-resolution spatial data in creating intuitive teleoperation systems. Both studies suggest that as the digital twin userbase grows, the user experience will become an essential factor, particularly in how objects are identified and interacted with within these virtual environments. Researchers should recognize that the usability of digital twins is as important as their technical accuracy [9].

Table 3. Examples of articles covering fundamental research in digital twins.

Domain	Physical Asset	Sensors	Physical–Digital Data Flow	Form of Digital Asset	Research Objective	Ref.
Sensor calibration	Building system and sensors	Temperature and pressure sensors	Sensor readings	Building automation system (BAS) interface	Simultaneous in situ calibration of physical and virtual sensors	[28]
Sensor reliability	Sensors	Vibration, temperature, and force sensors	Sensor readings	Simulated readings of the redundant digital sensors	Predict sensor failure using redundant digital sensors	[29]

Table 3. Cont.

Domain	Physical Asset	Sensors	Physical–Digital Data Flow	Form of Digital Asset	Research Objective	Ref.
Sensor connectivity	Human	Wearable ECG sensor	Sensor readings	Low-latency signal visualization and analysis	Enhanced biosensor data flow across a Bluetooth to Ethernet gateway	[30]
Tactile sensing	Tactile device	Tactile sensors	Touch events	Across-platform semantic abstraction and visualization	Ontology model construction for tactile sensing devices	[31]
Data processing	Truss	Force sensor and strain gauge	Sensor readings	Rapid visualization of the truss stress field	Multi-fidelity surrogate model development	[33]
Scene construction	Bridge construction site	Cameras	Multi-resolution images	Holographic scene in a prototype interface	Dynamic holographic modeling approach for augmented visualization of digital twin scenes	[34]
User interface design	Geometrical objects	Onboard robot sensing (e.g., LiDAR)	Point cloud	Visualization of the physical asset	Evaluate the recognition of physical objects and its implications on UI design for teleoperation systems	[32]

Table 4 presents a selection of research efforts showcasing the application of digital twin technology in the manufacturing and supply chain areas. In CNC machining, force sensors monitor the cutting torque in end milling processes, supplying data for a comprehensive dashboard that integrates real and simulated torque signals for condition monitoring [35]. The predictive maintenance capacity enables real-time adjustments and machine downtime reduction. Another study develops nonlinear multi-variant dynamic models of multi-axis machine tools with onboard CNC sensing data and visualizes the servo system’s dynamics [36]. The digital twins in the form of real-time visualization can help optimize the machine tool performance and reduce production errors.

Table 4. A sample of articles covering manufacturing and supply chain digital twin research.

Domain	Physical Asset	Sensors	Physical–Digital Data Flow	Form of Digital Asset	Research Objective	Ref.
CNC machining	Cutting torque in end milling	Force sensors	Process parameters and database of historical torque signals and analysis	Dashboard of simulated and real cutting torque and analysis	Machine tool condition monitoring	[35]
CNC machining	Servo system of a 5-axis laser drill	Onboard CNC sensing	In-process CNC data (e.g., body motion, actuator ripples, and vibration modes)	Real-time visualization of servo dynamic models	Develop nonlinear multi-variant dynamic models of multi-axis machine tools	[36]
CNC machining	Milling machine	Onboard CNC sensing (e.g., torque current and tachometer)	Fusion of tool, workpiece, and process monitoring data	Visual dashboard of part geometry, process data, and analysis	Development of digital process twin	[37]
Cyber-Physical System	CTTP 4.0 production cell	Optical sensors	Fusion of real-time operational parameters and sensor readings	Interactive visual replication of production cell	Practical implementation of digital twin complied to industry standards	[38]
Additive manufacturing	Temperature and strain profiles	Embedded distributed fiber sensors	Fusion of process parameters and sensor readings	FEA simulation of temperature and strain	Model temperature and strain with embedded distributed fiber sensors	[39]
Production planning	Body-in-white (BIW) production system	Onboard sensors and LiDAR	Fusion of CPS indicators, production data, and point cloud of the plant	3D production plant model with optimized production planning	Demonstrate the automated creation and updating of a BIW production digital twin	[40]

Further, in CNC machining, the fusion of tool, workpiece, and process monitoring data, is visualized on a dashboard, providing a complete view of the manufacturing process [37]. This digital process twin supports operators in making informed decisions by simulating part geometry and process analytics. The use of optical sensors in a cyber-physical production cell to create an interactive visual replica [38] signifies the importance of high-fidelity models for understanding and optimizing complex production systems. The above approaches to building digital twin models have made significant progress in unveiling the relations between key indicators and tool performance in the machining process. The sensor-based digital twins allow autonomous monitoring and troubleshooting within smart manufacturing environments. Future work could investigate the scalability of the method to consistently deliver accurate responses and optimize processes as the numbers of machine types and operational parameters scale. Additionally, exploring the integration of machine learning across different manufacturing environments would be valuable.

In additive manufacturing, embedded distributed fiber sensors are used for Finite Element Analysis (FEA) simulations of temperature and strain [39]. The ability to model these parameters with high precision is indicative of the move toward high-fidelity simulations in digital twins, ensuring product quality and process reliability.

Lastly, the production planning process benefits from the fusion of CPS indicators, production data, and LiDAR-generated point clouds to create a 3D model of a production plant [40]. This example demonstrates the potential of digital twins in providing a comprehensive three-dimensional context for production planning, facilitating better spatial understanding and resource allocation.

Table 5 presents the selected applications of digital twin sensor technology in the energy and power grid sector. The energy equipment monitoring example showcases a condition monitoring digital twin of a small hydro turbine, enabled by a wireless sensor network, including accelerometers, temperature, and inductive current sensors. The digital twins operating on sensor readings and environmental data provide a condition indicator visualization [41]. This approach can detect faults early and reduce downtime, which is crucial in the energy sector where continuity is essential.

Table 5. A selection of research articles covering energy and power grid digital twins research.

Domain	Physical Asset	Sensors	Physical–Digital Data Flow	Form of Digital Asset	Research Objective	Ref.
Equipment monitoring	Small hydro turbine	Wireless sensor network (e.g., accelerometers, temperature, and inductive current sensors)	Sensor readings, production, and environmental data	Condition indicator visualization and predictive analysis	Build digital twin from various data source for turbine condition monitoring	[41]
Electric power converter	Photovoltaic (PV) dc–dc converter	Thermal camera	Scanning electron microscope (SEM) images, converter tech specs, and temperature data	FEM simulation and prediction of junction, case, and heat sink temperatures	Fast estimation of switching device temperature in PV converters	[42]
Wind engineering	Wind pressure field	Wind pressure sensors	Sensor readings	Wind pressure field visualization and measurements	Develop an optimal sensor placement algorithm to reconstruct wind pressure fields	[43]
Hydropower generation	Hydraulic network (e.g., valve, tank, and pipe)	Pressure sensors	Hydraulic model of the plant and the valve and flap configuration	Hydraulic variables and hydropower generation visualization and analysis	Develop a control system to maximize hydropower production while meeting hydraulic constraints	[44]
Smart grid	Smart grid (e.g., consumer, producer, and regulator)	Event logger	Live semantic annotation (LSA) of events and coordination laws that cause the events to evolve	Autonomous proactive agents on a coordination platform CLEMAP	Develop a series of digital twins that interact and coordinate activities to exchange energy and enhance grid stability	[45]

In electric power conversion, the photovoltaic (PV) dc–dc converter’s efficiency is augmented by thermal cameras and scanning electron microscope imagery. FEM simulations predict temperatures at critical converter components, enabling fast estimations of device conditions under various operational stresses [42]. The two studies highlight the importance of digital twin predictive maintenance in infrastructure reliability.

Wind engineering research utilizes wind pressure sensors to develop an optimal sensor placement algorithm [43]. This algorithm aims to reconstruct wind pressure fields accurately, which is indispensable for assessing the structural integrity of wind turbines and optimizing their design for maximum energy capture. The reconstruction of the wind pressure field can also be utilized for creating digital twin infrastructure.

For hydropower generation, pressure sensors are deployed within the hydraulic network, informing the development of a control system that maximizes hydropower production while adhering to hydraulic constraints [44]. Such digital twin applications ensure the harmonization of power generation with environmental and infrastructural considerations.

Lastly, in the field of smart grids, event loggers are utilized to develop digital twins with autonomous proactive agents [45]. These agents interact within a coordination platform to manage the complex dynamics of energy demand and supply, thus enhancing grid stability and operational resilience. This research points out that integrating an agent-coordination model into digital twins can address complex energy management issues at the microgrid level. The findings provide an example of how to create resilient and user-centric energy networks.

Table 6 presents the applications of digital twin sensor technology in the healthcare and human-centric area, where sensors are broadly referred to as any device or system that detects events or changes in a given environment, transmitting the information to other devices. In the cardiology field, ECG sensors are integral in developing a digital twin of the human heart [46]. This innovative approach merges ECG data with medical records to construct a “Cardio Twin”, a proof of concept that offers heart condition visualizations for both local and remote diagnosis. In another example, clinical 12-lead ECGs and Magnetic Resonance Imaging (MRI) create biophysically detailed digital twins for cardiac electrophysiology [47]. These models simulate intricate heart structures, including Purkinje networks, paving the way for in silico clinical trials and advanced cardiac care.

Table 6. A selection of research articles covering healthcare and human-centric digital twin research.

Domain	Physical Asset	Sensors	Physical–Digital Data Flow	Form of Digital Asset	Research Objective	Ref.
Cardiology	Human Heart	ECG sensors	ECG data and medical records	Heart condition visualization and analysis available for local and remote diagnosis	Proof of concept version of the Cardio Twin	[46]
Cardiology	Human heart and cardiac electrophysiology	Clinical 12-lead ECG and magnetic resonance (CMR) imaging	ECG data and CMR imaging	Biophysically detailed cardiac twin including Purkinje networks and cardiac electrophysiology	Create personalized, multiscale biventricular heart models for in silico clinical trial	[47]
Rural health	Patients in rural areas	Healthcare IoT sensors and medical devices	Sensor readings	Blockchain-encrypted medical data and analysis	Integrate healthcare IoT data with blockchain for secure and efficient data management in rural healthcare	[48]
Space medicine	Medical training environment	Mixed reality (e.g., HoloLens) and haptic devices	Digitized real-world training scenarios and learner’s real-time input	Interactive training with virtual feedback integrated into real-world scenarios	Develop a mixed-reality-based medical training platform for astronauts	[49]

Table 6. Cont.

Domain	Physical Asset	Sensors	Physical–Digital Data Flow	Form of Digital Asset	Research Objective	Ref.
Education	Remote lab of a production cell	Onboard sensors for equipment control and monitoring	Sensor readings	Interactive interface with multimodal visualizations	Apply the digital twin concept in a hybrid remote laboratory for various learning scenarios	[50]
Human–robot collaboration	Battery pack assembly line	Force/torque sensor	Fusion of sensor readings and production data (e.g., idle times)	Visualization and analysis of the human–robot collaboration assembly line	Design, develop, and operate an agile, adaptable, and safe human–robot collaborative system	[51]

For rural healthcare, IoT sensors and devices are leveraged to bring medical services to remote areas [48]. Here, sensors encompass a variety of medical devices that collect health-related data, which, when coupled with blockchain technology, ensures secure data management and analysis in resource-limited settings. In space medicine, sensors include mixed reality devices such as HoloLens and haptic systems, which create a digitized interactive training environment [49]. This expands the sensory experience by providing real-time feedback and immersive scenarios for astronaut medical training. The above research highlights the potential and success of digital twin technology in the field of personalized and predictive healthcare.

In the educational sector, sensors refer to the instrumentation of a remote lab, where equipment control and monitoring are critical [50]. These sensor systems enable a hybrid remote laboratory for various learning scenarios, fostering interactive and multimodal educational experiences. This also encourages future researchers to explore digital twin solutions for better learning outcomes and operational safety [52].

Lastly, in the context of human–robot collaboration, force/torque sensors on a battery pack assembly line provide data for a digital twin that visualizes and analyzes the collaborative environment [51]. This digital twin assists in designing, developing, and operating a safe and efficient human–robot interactive system. Future study can revolve around the safety and optimization of these systems with the aid of digital twins [53].

Table 7 presents examples of the use of digital twins in the optimization of infrastructure and transportation systems. For infrastructure modeling, LiDAR sensors are utilized to capture detailed point cloud data of campus buildings [54]. This technology enables the creation of accurate digital replicas of large structures and facilitates the creation of accurate digital twins of extensive structures, enabling efficient maintenance planning and historical preservation.

In transportation infrastructure, the fusion of 2D images from cameras and 3D point clouds from LiDAR leads to a comprehensive digital twin of a magnetic levitation track [55]. This detailed representation bridges the gap between macroscopic project management and microscopic engineering analysis, underscoring the capacity for digital twins to offer multiscale insights into transportation systems. This study points out the importance of having efficient and automated processes for managing large LiDAR datasets to enhance the scalability of digital twins in civil engineering. Future study should include developing advanced algorithms that could automate the conversion of point cloud data into information modeling.

Urban logistics can benefit from the integration of sensors and actuators within the infrastructure. This can be achieved through a platform architecture for digital twins that informs policy making via interactive dashboards [56]. This approach allows for real-time sensor data and logistics system documentation to drive simulation models that can pinpoint gaps and opportunities for transformation within city ecosystems. It is challenging to convert this framework into physical systems for city planners and logistics stakeholders to use and improve urban logistics.

Table 7. A selection of infrastructure and transportation digital twins research.

Domain	Physical Asset	Sensors	Physical–Digital Data Flow	Form of Digital Asset	Research Objective	Ref.
Infrastructure modeling	Campus buildings	LiDAR	Point cloud data	Virtual replicas of large campus infrastructure	Assess and implement reconstruction methods to create digital twins of large infrastructure using point cloud data	[54]
Transportation infrastructure	Magnetic levitation track	Cameras and LiDAR	Fusion of 2D images 3D point clouds	Detailed digital representation with macroscopic to microscopic perspective analysis	Fuse 2D image and 3D point cloud to create a digital twin model of magnetic levitation track	[55]
Urban logistics	Urban infrastructure and logistics systems	Sensors, actuators, and logistic system documentation	Sensor readings reflecting the state of the logistic system	Interactive dashboards with metrics and logs to inform policy making	Propose a platform architecture for digital twins in urban logistics, addressing gaps in simulation model orchestrations and data transformation	[56]
Factory logistics	Assembly line with Automated Guided Vehicles (AGVs)	Sensors for AGV tracking and monitoring	Sensor readings and documentation of assembly logistics	Data model, simulation analysis, and virtual action model of the assembly line	Develop and apply an AGV multi-objective dynamic scheduling method based on digital twins to improve logistics efficiency	[34]
Transportation planning	Long-distance freight flows with various transport modes	IoT sensors, GPS, and GIS	Real-time data inputs from transport modes	Virtual infrastructure visualization and transportation mode analysis	Explore the potential of digital twins in synchromodal transport for long-distance freight flows	[57]
Smart agriculture	Cyclone bag filter in grain milling plants	Pressure sensors and anemometer	Real-time sensor readings	Computational fluid dynamics simulation of cyclone bag filter for RUL prediction	Apply digital twins for predictive maintenance of the cyclone bag filter system	[58]

Factory logistics can be revolutionized by incorporating Automated Guided Vehicles (AGVs) that track and monitor the movement of goods on the assembly line [34]. The development and application of a multi-objective AGV scheduling method based on digital twins reflect a shift toward intelligent and efficient logistics systems.

The planning of long-distance freight flows can be analyzed by integrating IoT sensors, GPS, and GIS into a virtual infrastructure and transportation model [57]. This digital twin serves as a powerful tool for analyzing and synchronizing transport, demonstrating the potential of digital twins to streamline logistics operations across vast distances. Future study can focus on achieving interconnection between real-time data and virtual models for different transport modes in an operational context.

The predictive maintenance of agriculture equipment can be improved by digital twin technologies integrated with sensors and data pipeline systems [58]. This study streamlines computational fluid dynamics (CFD) simulation data, sensor readings, and historical information to replicate a virtual cyclone bag filter system in grain milling plants. The digital twins of the system can monitor the filter status and perform precise predictions of a system’s remaining useful life. This research marks the potential of digital twins in improving operational efficiency for smart agriculture through monitoring and predictive analytics.

Table 8 explores the digital twin applications in business and asset management. These sensors are not limited to traditional physical devices but also include digital and social data sources. In production management, the term “sensor” encompasses product documentation throughout production [59]. This documentation acts as a sensor by providing continuous feedback on the product lifecycle, enabling the development of a digital twin for efficient tracking in high-volume production environments. This study set an example

of using Asset Administration Shell to standardize and simplify the digital twin representations of manufacturing assets. Another study in this domain proposes a hybrid digital twin approach that integrates traditional onboard sensors with telemetry data sources to create virtual production line properties [60]. Their innovative usage of Apache StreamPipes for handling high-volume data streams features a solution to the data preprocessing for digital twins.

Table 8. A selection of business and asset management digital twins research.

Domain	Physical Asset	Sensors	Physical–Digital Data Flow	Form of Digital Asset	Research Objective	Ref.
Production management	Products and assets	Product documentation throughout production	Data from production assets, plants, and operators	Asset Administration Shell (AAS) for lifecycle tracking	Develop a digital twin for efficient lifecycle tracking in high-volume production	[59]
Production management	Steel production line	Onboard sensors and telemetry data sources in production line	Sensor readings, batch data, and manually-input data	Virtual production line properties via the digital twin API	Propose a hybrid digital twin approach integrating Industry 4.0-compliant digital twins with AAS and Apache StreamPipes	[60]
Environment monitoring	Spread of invasive species	Social Media (Twitter)	Posts on spotted lanternfly sightings and behaviors	Visualization of invasive species spread	Create a digital twin using social media content for environmental monitoring	[61]
Social issue alleviation	Social subjects (e.g., personal life, organizations)	Chat rooms	Communication data from chat rooms	Real-time sentiment analysis and Chatbot-driven conversation facilitation	Use chatbots to stimulate and sustain communication in social environments	[62]
Social network	Social sentiments	Business intelligence data infrastructure	Data from social networks	Knowledge graphs for social data mapping	Create semantic-driven digital twins for analyzing social network dynamics	[63]

The notion of sensors expands further in environmental monitoring, where social media posts on platforms such as Twitter become inputs. These “digital sensors” capture real-time data on the spread of invasive species [61], offering a novel approach to environmental monitoring by harnessing crowd-sourced information. This study manifests the versatile nature of digital twins by bridging it with ecological management and leveraging Natural Language Processing to model the spread of an invasive species. Additionally, in social issue alleviation and network sentiment analysis, chat rooms and business intelligence data act as sensors by providing communication data and social network dynamics [62,63]. These data allow for real-time sentiment analysis and conversation facilitation via chatbots. The concept of semantic digital twins for simulating human behavior for analytical purposes is an innovative idea. In future studies, it is important to address privacy and data security concerns, such as modeling complex human behavior and ensuring ethical use of personal data.

4. Discussion

Our analysis of digital twin research from 2000 to 2023 shows that the field has been growing and diversifying rapidly. This trend, particularly notable post-2020, offers digital twin researchers avenues for new research opportunities and gives digital twin architects insights into the evolving applications in areas such as smart cities, healthcare, and energy. Emerging functions such as decision making and predictive maintenance also demonstrate the field’s advancement.

Sensor-related keywords, such as real time, point cloud, and sensor network, are becoming more important. This emphasizes the demand for real-time, interconnected, and multimodal sensor data. This trend aligns with the computational architecture shift from cloud to edge computing, indicating a move toward distributed computing for faster and

more efficient data processing. Moreover, the emergence of advanced machine learning models such as deep learning and federated learning reflects the increasing complexity of sensor data processing. It also highlights the growing importance of privacy and security for digital twins.

According to the keyword trends, digital twins are advancing beyond conventional simulations such as FEA. The research suggests that there is a shift toward developing systems that not only replicate physical entities but also evolve with them. The importance of keywords such as human–robot interaction and predictive maintenance has been growing, which indicates the emergence of interactive and preemptive digital twins.

The Sankey diagrams demonstrate the wide range of sensor technologies used in digital twins. These sensors go beyond traditional physical devices and include medical equipment, clinical tests, and even social media platforms. This approach allows digital twins to provide a comprehensive and accurate representation of real-world scenarios. Moreover, the literature highlights the rise of virtual and soft sensors, which mirror physical sensor functions, offering preemptive insights and facilitating proactive sensing system maintenance.

Our study also revealed a connection between sensor selection and the functionality levels of digital twins, namely representation, replication, reality, and relational. From the review of specific digital twin examples, we observe that the choice of sensors directly influences the level of digital twin functionality, particularly in retrofit designs. As the digital twin capacity increases, we also observe that many applications need sensor data fusion. For example, in the manufacturing setting, a proactive machine tool digital twins will need the input of low-frequency production data and high-frequency sensor data. Incorrect sensor choices can lead to a mismatch between sensor capacity and the expected functionality of digital twins, eventually impacting the performance of the digital twins.

In addition, we believe that relying heavily on data-driven insights without substantial domain expertise can be risky. With limited domain knowledge, there is a possibility of creating brittle and underwhelming digital twins that may not respond inadequately to real-world variables. Therefore, future research should aim at integrating domain knowledge and robust empirical knowledge into digital twins to enhance their reliability and accuracy.

5. Conclusions

This study analyzed the research trends in the field of digital twins by examining metadata from 9639 peer-reviewed articles published between 2000 and 2023. We processed the metadata using an NLP-based toolkit and manually labeled each article with its most relevant application field. Using the KCN methodology, we performed temporal research trend analysis, mapping popular sensing technologies to six application fields and identifying representative examples of digital twins in each field. For researchers, this analysis provides a comprehensive view of the field's development, identifying key areas for future exploration. For architects, the findings highlight technological applications and examples essential for informed decision making in digital twin system design.

This study found that the field of digital twins is rapidly growing and diversifying. We used network metrics to analyze the temporal changes in the field and identified emerging and declining keywords over time. We also identified emerging application fields, functions, and enabling sensing technologies. The findings suggest that digital twins are moving toward predictive tasks while ensuring system integrity and security across many sectors beyond manufacturing.

We used a Sankey chart to visualize the mapping from popular sensing technologies to six application fields. We found that real-time data, point cloud data, and human–robot interaction are increasing trends. Additionally, we noticed an extension of the traditional sensor definition to include novel sensors such as medical tests and social media posts. We identified neural networks and reinforcement learning as crucial for autonomous decision making. The emergence of federated learning marks a shift toward distributed computation, emphasizing data privacy.

Following the mapping, we reviewed specific examples of digital twins in each field. For each application, we analyzed its physical assets, sensors, physical–digital data flow, the form of digital assets, and research objectives. From these examples, we observed a connection between sensor selection and the functionality level of digital twins. We raised concerns over the mismatch between sensor capacity and digital twins’ functionality and possible brittle digital twins if they are too dependent on empirical prior knowledge.

Author Contributions: Conceptualization, S.K. and W.L.; methodology, S.K., W.L. and Z.L.; software, Z.L. and W.L.; validation, H.Z.; formal analysis, W.L. and H.Z.; investigation, S.K. and W.L.; resources, S.K.; data curation, W.L., Z.L. and H.Z.; writing—original draft preparation, W.L. and H.Z.; writing—review and editing, S.K. and Z.L.; visualization, W.L.; supervision, S.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Schwer, L.E. An overview of the ASME V&V-10 guide for verification and validation in computational solid mechanics. In Proceedings of the 20th International Conference on Structural Mechanics in Reactor Technology, Espoo, Finland, 9–14 August 2009; pp. 1–10.
2. Grieves, M. Digital twin: Manufacturing excellence through virtual factory replication. *White Pap.* **2014**, *1*, 1–7.
3. Van der Valk, H.; Haße, H.; Möller, F.; Arbter, M.; Henning, J.L.; Otto, B. A Taxonomy of Digital Twins. In Proceedings of the AMCIS, Online, 15–17 August 2020.
4. Wooley, A.; Silva, D.F.; Bitencourt, J. When is a simulation a digital twin? A systematic literature review. *Manuf. Lett.* **2023**, *35*, 940–951. [[CrossRef](#)]
5. Lysova, N.; Solari, F.; Vignali, G. Optimization of an indirect heating process for food fluids through the combined use of CFD and Response Surface Methodology. *Food Bioprod. Process.* **2022**, *131*, 60–76. [[CrossRef](#)]
6. Atzori, L.; Iera, A.; Morabito, G. The internet of things: A survey. *Comput. Netw.* **2010**, *54*, 2787–2805. [[CrossRef](#)]
7. Lee, E.A. The past, present and future of cyber-physical systems: A focus on models. *Sensors* **2015**, *15*, 4837–4869. [[CrossRef](#)] [[PubMed](#)]
8. Tao, F.; Qi, Q.; Wang, L.; Nee, A. Digital twins and cyber-physical systems toward smart manufacturing and industry 4.0: Correlation and comparison. *Engineering* **2019**, *5*, 653–661. [[CrossRef](#)]
9. Minerva, R.; Lee, G.M.; Crespi, N. Digital twin in the IoT context: A survey on technical features, scenarios, and architectural models. *Proc. IEEE* **2020**, *108*, 1785–1824. [[CrossRef](#)]
10. American Industrial Hygiene Association. *The Future of Sensors—Protecting Worker Health Through Sensor Technologies*; American Industrial Hygiene Association (AIHA): Falls Church, VA, USA, 2016.
11. Radhakrishnan, S.; Erbis, S.; Isaacs, J.A.; Kamarthi, S. Novel keyword co-occurrence network-based methods to foster systematic reviews of scientific literature. *PLoS ONE* **2017**, *12*, e0172778.
12. Liu, M.; Fang, S.; Dong, H.; Xu, C. Review of digital twin about concepts, technologies, and industrial applications. *J. Manuf. Syst.* **2021**, *58*, 346–361. [[CrossRef](#)]
13. Botín-Sanabria, D.M.; Mihaita, A.S.; Peimbert-García, R.E.; Ramírez-Moreno, M.A.; Ramírez-Mendoza, R.A.; Lozoya-Santos, J.d.J. Digital Twin Technology Challenges and Applications: A Comprehensive Review. *Remote Sens.* **2022**, *14*, 1335. [[CrossRef](#)]
14. Mihai, S.; Yaqoob, M.; Hung, D.V.; Davis, W.; Towakel, P.; Raza, M.; Karamanoglu, M.; Barn, B.; Shetve, D.; Prasad, R.V.; et al. Digital Twins: A Survey on Enabling Technologies, Challenges, Trends and Future Prospects. *IEEE Commun. Surv. Tutor.* **2022**, *24*, 2255–2291. [[CrossRef](#)]
15. Sharma, A.; Kosasih, E.; Zhang, J.; Brintrup, A.; Calinescu, A. Digital Twins: State of the art theory and practice, challenges, and open research questions. *J. Ind. Inf. Integr.* **2022**, *30*, 100383. [[CrossRef](#)]
16. Leng, J.; Wang, D.; Shen, W.; Li, X.; Liu, Q.; Chen, X. Digital twins-based smart manufacturing system design in Industry 4.0: A review. *J. Manuf. Syst.* **2021**, *60*, 119–137. [[CrossRef](#)]
17. van der Valk, H.; Strobel, G.; Winkelmann, S.; Hunker, J.; Tomczyk, M. Supply Chains in the Era of Digital Twins—A Review. *Procedia Comput. Sci.* **2022**, *204*, 156–163. [[CrossRef](#)]
18. Bado, M.F.; Tonelli, D.; Poli, F.; Zonta, D.; Casas, J.R. Digital Twin for Civil Engineering Systems: An Exploratory Review for Distributed Sensing Updating. *Sensors* **2022**, *22*, 3168. [[CrossRef](#)] [[PubMed](#)]
19. Jafari, M.; Kavousi-Fard, A.; Chen, T.; Karimi, M. A Review on Digital Twin Technology in Smart Grid, Transportation System and Smart City: Challenges and Future. *IEEE Access* **2023**, *11*, 17471–17484. [[CrossRef](#)]
20. Pylianidis, C.; Osinga, S.; Athanasiadis, I.N. Introducing digital twins to agriculture. *Comput. Electron. Agric.* **2021**, *184*, 105942. [[CrossRef](#)]
21. Purcell, W.; Neubauer, T. Digital Twins in Agriculture: A State-of-the-art review. *Smart Agric. Technol.* **2023**, *3*, 100094. [[CrossRef](#)]

22. Preite, L.; Solari, F.; Vignali, G. Technologies to Optimize the Water Consumption in Agriculture: A Systematic Review. *Sustainability* **2023**, *15*, 5975. [\[CrossRef\]](#)
23. Volkov, I.; Radchenko, G.; Tchernykh, A. Digital Twins, Internet of Things and Mobile Medicine: A Review of Current Platforms to Support Smart Healthcare. *Program. Comput. Softw.* **2021**, *47*, 578–590. [\[CrossRef\]](#)
24. Armeni, P.; Polat, I.; De Rossi, L.M.; Diaferia, L.; Meregalli, S.; Gatti, A. Digital Twins in Healthcare: Is It the Beginning of a New Era of Evidence-Based Medicine? A Critical Review. *J. Pers. Med.* **2022**, *12*, 1255. [\[CrossRef\]](#)
25. Wuttke, H.D.; Henke, K.; Hutschenreuter, R. Digital Twins in Remote Labs. In Proceedings of the Cyber-Physical Systems and Digital Twins, Bangalore, India, 3–6 February 2019; Auer, M.E., Kalyan Ram B., Eds.; Lecture Notes in Networks and Systems; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 289–297.
26. Ozek, B.; Lu, Z.; Pouromran, F.; Radhakrishnan, S.; Kamarthi, S. Analysis of pain research literature through keyword Co-occurrence networks. *PLoS Digit. Health* **2023**, *2*, e0000331. [\[CrossRef\]](#)
27. Wang, M.; Chai, L. Three new bibliometric indicators/approaches derived from keyword analysis. *Scientometrics* **2018**, *116*, 721–750. [\[CrossRef\]](#)
28. Koo, J.; Yoon, S. Simultaneous in-situ calibration for physical and virtual sensors towards digital twin-enabled building operations. *Adv. Eng. Inform.* **2024**, *59*, 102239. [\[CrossRef\]](#)
29. Fett, M.; Turner, E.; Breimann, R.; Kirchner, E. Extension of the system boundary of the Digital Twin onto the sensors of the Physical Twin through the introduction of redundant soft sensors. *Forsch. Ingenieurwesen* **2023**, *87*, 479–488. [\[CrossRef\]](#)
30. Grammatikakis, M.; Ninidakis, S.; Kornaros, G.; Bakoyiannis, D. Towards Efficient Gateways and Servers for Biosensors. In Proceedings of the International Conference on Applications in Electronics Pervading Industry, Environment and Society; Springer: Berlin/Heidelberg, Germany, 2022; pp. 346–352.
31. Van Nguyen, P.; Pham, V.C.; Tan, Y. Toward a tactile ontology for semantic interoperability of the tactile internet. In Proceedings of the 2022 IEEE 16th International Conference on Semantic Computing (ICSC), Laguna Hills, CA, USA, 26–28 January 2022; pp. 115–118.
32. Bremner, P.; Giuliani, M. Impact of resolution, colour, and motion on object identification in digital twins from robot sensor data. *Front. Robot.* **2022**, *9*, 995342. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Wang, S.; Lai, X.; He, X.; Qiu, Y.; Song, X. Building a trustworthy product-level shape-performance integrated digital twin with multifidelity surrogate model. *J. Mech. Des.* **2022**, *144*, 031703. [\[CrossRef\]](#)
34. Wu, J.; Zhu, J.; Zhang, J.; Dang, P.; Li, W.; Guo, Y.; Fu, L.; Lai, J.; You, J.; Xie, Y.; et al. A dynamic holographic modelling method of digital twin scenes for bridge construction. *Int. J. Digit. Earth* **2023**, *16*, 2404–2425. [\[CrossRef\]](#)
35. Ghosh, A.K.; Ullah, A.S.; Teti, R.; Kubo, A. Developing sensor signal-based digital twins for intelligent machine tools. *J. Ind. Inf. Integr.* **2021**, *24*, 100242. [\[CrossRef\]](#)
36. Wang, C.P.; Erkorkmaz, K.; McPhee, J.; Engin, S. In-process digital twin estimation for high-performance machine tools with coupled multibody dynamics. *CIRP Ann.* **2020**, *69*, 321–324. [\[CrossRef\]](#)
37. Caesar, B.; Hänel, A.; Wenkler, E.; Corinth, C.; Ihlenfeldt, S.; Fay, A. Information model of a digital process twin for machining processes. In Proceedings of the 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vienna, Austria, 8–11 September 2020; Volume 1, pp. 1765–1772.
38. Krenczyk, D. Digital Twins of the Manufacturing System Based on a Simulation Model with Bi-directional Process Data Exchange. In Proceedings of the International Conference on Intelligent Systems in Production Engineering and Maintenance, Wrocław, Poland, 13–15 September 2023; Springer: Berlin/Heidelberg, Germany, 2023; pp. 548–559.
39. Zou, R.; Liang, X.; Chen, Q.; Wang, M.; Zaghoul, M.A.; Lan, H.; Buric, M.P.; Ohodnicki, P.R.; Chorpene, B.; To, A.C.; et al. A digital twin approach to study additive manufacturing processing using embedded optical fiber sensors and numerical modeling. *J. Light. Technol.* **2020**, *38*, 6402–6411. [\[CrossRef\]](#)
40. Biesinger, F.; Meike, D.; Kraß, B.; Weyrich, M. A case study for a digital twin of body-in-white production systems general concept for automated updating of planning projects in the digital factory. In Proceedings of the 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), Turin, Italy, 4–7 September 2018; Volume 1, pp. 19–26.
41. Ravet, F.; Nicoulaz, D.; Jacquat, Y.; Morier, C.; André, N. Wireless Sensor Network Powered with Data Analytics for Small Hydro Reliability Enhancement. In Proceedings of the 2023 IEEE Sensors, Vienna, Austria, 29 October–1 November 2023; pp. 1–4.
42. Van Cappellen, L.; Deckers, M.; Alavi, O.; Daenen, M.; Driesen, J. A Real-time Physics Based Digital Twin for Online MOSFET Condition Monitoring in PV Converter Applications. In Proceedings of the 2022 28th International Workshop on Thermal Investigations of ICs and Systems (THERMINIC), Dublin, Ireland, 28–30 September 2022; pp. 1–4.
43. Luo, X.; Kareem, A.; Yoo, S. Optimal sensor placement for reconstructing wind pressure field around buildings using compressed sensing. *J. Build. Eng.* **2023**, *75*, 106855. [\[CrossRef\]](#)
44. Sinagra, M.; Creaco, E.; Morreale, G.; Tucciarelli, T. Energy recovery optimization by means of a turbine in a pressure regulation node of a real water network through a data-driven digital twin. *Water Resour. Manag.* **2023**, *37*, 4733–4749. [\[CrossRef\]](#)
45. Glass, P.; Di Marzo Serugendo, G. Coordination Model and Digital Twins for Managing Energy Consumption and Production in a Smart Grid. *Energies* **2023**, *16*, 7629. [\[CrossRef\]](#)
46. Martinez-Velazquez, R.; Gamez, R.; El Saddik, A. Cardio Twin: A Digital Twin of the human heart running on the edge. In Proceedings of the 2019 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Istanbul, Turkey, 26–28 June 2019; pp. 1–6.

47. Camps, J.; Berg, L.A.; Wang, Z.J.; Sebastian, R.; Riebel, L.L.; Doste, R.; Zhou, X.; Sachetto, R.; Coleman, J.; Lawson, B.; et al. Digital Twinning of the Human Ventricular Activation Sequence to Clinical 12-lead ECGs and Magnetic Resonance Imaging Using Realistic Purkinje Networks for in Silico Clinical Trials. *arXiv* **2023** arXiv:2306.13740.
48. Kerrison, S.; Jusak, J.; Huang, T. Blockchain-Enabled IoT for Rural Healthcare: Hybrid-Channel Communication with Digital Twinning. *Electronics* **2023**, *12*, 2128. [[CrossRef](#)]
49. Ip-Jewell, S.; Jewell, E.H. Holotriage: A Novel Medical First Response Training for Astronauts Integrating Artificial Intelligence, Digital Twins, Avatars, Haptics, and Mixed Reality Spatial Computing Technologies. In Proceedings of the 73rd International Astronautical Congress (IAC-22), International Astronautical Federation (IAF), Paris, France, 18–22 September 2022.
50. Wuttke, H.D.; Henke, K.; Hutschenreuter, R. Digital twins in remote labs. In *Cyber-Physical Systems and Digital Twins, Proceedings of the 16th International Conference on Remote Engineering and Virtual Instrumentation 16, Bangalore, India, 3–6 February 2019*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 289–297.
51. Malik, A.A.; Brem, A. Digital twins for collaborative robots: A case study in human-robot interaction. *Robot. Comput.-Integr. Manuf.* **2021**, *68*, 102092. [[CrossRef](#)]
52. Corter, J.E.; Esche, S.K.; Chassapis, C.; Ma, J.; Nickerson, J.V. Process and learning outcomes from remotely-operated, simulated, and hands-on student laboratories. *Comput. Educ.* **2011**, *57*, 2054–2067. [[CrossRef](#)]
53. Schröder, I.; Huang, D.Y.Q.; Ellis, O.; Gibson, J.H.; Wayne, N.L. Laboratory safety attitudes and practices: A comparison of academic, government, and industry researchers. *J. Chem. Health Saf.* **2016**, *23*, 12–23. [[CrossRef](#)]
54. Luis Lugo, J.; Gallegos Reyes, J.C.; Weidner, J.; Raheem, A.; Cheu, R.L. Evaluation of 3D Reconstruction Methods from Terrestrial LiDAR Point Cloud Data to Create Digital Twins of Civil Infrastructure Projects. In Proceedings of the International Conference on Transportation and Development 2022, Washington, DC, USA, 31 May–3 June 2022; pp. 81–92.
55. Zhang, Y.; Zhang, L.; Shen, G. Multi-Source and Multi-Dimensional Data Fusion of Magnetic Levitation Track Transportation Based on Digital Twin. In Proceedings of the 2022 International Conference on Computer Science, Information Engineering and Digital Economy (CSIEDE 2022), Guangzhou, China, 28–30 October 2022; Atlantis Press: Dordrecht, The Netherlands, 2022; pp. 595–610.
56. Belfadel, A.; Hörl, S.; Tapia, R.J.; Politaki, D.; Kureshi, I.; Tavasszy, L.; Puchinger, J. A conceptual digital twin framework for city logistics. *Comput. Environ. Urban Syst.* **2023**, *103*, 101989. [[CrossRef](#)]
57. Ambra, T.; Macharis, C. Agent-based digital twins (ABM-DT) in synchromodal transport and logistics: The fusion of virtual and physical spaces. In Proceedings of the 2020 Winter Simulation Conference (WSC), Orlando, FL, USA, 14–18 December 2020; pp. 159–169.
58. Solari, F.; Lysova, N.; Montanari, R. Digital Twin Based on Historical Data and Simulation Results: Fault Detection and Estimation of the Remaining Useful Life of a Cyclone Bag Filter. *Appl. Sci.* **2023**, *13*, 8297. [[CrossRef](#)]
59. Gärtner, S.; Brandstetter, A.; Oberle, M. A Blueprint Of Digital Twins In High Volume Production Environments Using The Asset Administration Shell. In Proceedings of the 2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA), Stuttgart, Germany, 6–9 September 2022; pp. 1–7.
60. Jacoby, M.; Jovicic, B.; Stojanovic, L.; Stojanović, N. An approach for realizing hybrid digital twins using asset administration shells and apache StreamPipes. *Information* **2021**, *12*, 217. [[CrossRef](#)]
61. Scheibmeir, J.A.; Malaiya, Y.K. A Social Media-driven Digital Twin of an Invasive Species. In Proceedings of the 10th International Workshop on Simulation for Energy, Sustainable Development & Environment, SESDE, Rome, Italy, 19–21 September 2022.
62. Wunderlich, W. Chat Bot Concept for a Social Pain Reliever. In Proceedings of the 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET), Cape Town, South Africa 9–10 December 2021; pp. 1–6.
63. Berlanga, R.; Museros, L.; Llidó, D.M.; Sanz, I.; Aramburu, M.J. Towards Semantic DigitalTwins for Social Networks. In Proceedings of the Second International Workshop on Semantic Digital Twins, Hersonissos, Greece, 6 June 2021.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.