INDUCTIVE THEORIZATION OF CONSTRUCTION DIGITAL TWINS: DERIVING A TAXONOMY OF APPLICATIONS AND ENABLING TECHNOLOGIES

Wahib Saif 1,2, SeyedReza RazaviAlavi1, Kay Rogage1, Xiang Xie3, and Mohamad Kassem3
1Northumbria University, Newcastle, UK
2BIM Academy, Newcastle, UK
3Newcastle University, Newcastle, UK

Abstract

Digital twin in the construction phase—termed Construction Digital Twin (CDT)—faces more significant challenges than in other phases, such as operation or maintained phase, due to the dynamic and evolving nature of construction sites and their broad spectrum of applications. To enhance our understanding of the CDT domain, it is crucial to clearly define it, establish a detailed taxonomy of its current applications and enabling technologies, and elucidate how attributes, data requirements, and technology choices within CDT vary across different construction applications. This objective remains unmet in existing literature, a gap this paper addresses through an approach combining systematic review, thematic coding, and conceptualization of CDT architecture comprising of five layers: sensing, communication, storage, analytics, and visualization. The study identifies seven major applications of CDT and maps them to the five architectural layers and their enabling technologies, providing insights into the suitability and prevalence of these technologies for specific applications of interest.

Introduction

Digital twin (DT) is an emerging and promising advancement that has the potential to transform the construction industry and drive its productivity (Boje et al., 2020; Opoku et al., 2021). DT is generally perceived as an up-to-date digital representation of a physical system and its functional properties leveraging data streaming from a wide range of data capturing and communication technologies (Sacks et al., 2020). It can provide data-driven insights about the state and performance of an object or a process in real time thus enabling informed decision-making capabilities. DT generally has three main elements: the physical environment, its digital representation, and the streaming of data/feedback that connects the two, enabled by data sensing technologies, artificial intelligence functions, and high-speed networking (Sacks et al., 2020).

Although DT is still regarded as a nascent concept within the construction industry, it has attracted substantial attention from researchers over the past five years investigating its applications during the different phases of built asset projects, with a particular focus on the design and operational phases (Opoku et al., 2021). DT applications during the construction phase have been emerging recently and still require extensive research. Hence, this paper aims to present a systematic approach to reviewing, mapping, and classifying the different applications of DT exclusively during the construction phase of built assets, namely Construction Digital Twin (CDT). Doing so provides transparency to readers on the most comprehensive state-of-the-art applications of DT and further supports its adoption during the construction phase. This paper also aims to thoroughly examine the different DT systems presented within the literature to give a generalized anatomy of DT architecture and the different technologies integrated within its identified layers. This will provide knowledge of the concepts, technologies, and processes required to replicate or build upon existing systems thus enhancing the scalability and adaptability of future-developed DTs.

Research Methodology

A systematic review was performed by following the reporting checklist of PRISMA to present a transparent, replicable, updateable, and comprehensive summary of previous studies related to CDT domain. Three databases were selected (i.e., Scopus, Web of Science, and Google Scholar) for carrying out the preliminary literature search. For Google Scholar, the search string included keywords of “Digital twin” and “Construction site” or “Construction Phase” and only results within the first 30 pages were included as results were not relevant anymore to the scope of this search. For Scopus and Web of Science, these keywords were expanded to form a broad search string presented in Table 1. The search, unrestricted by publication date up to July 2023, retrieved 504 DT publications from three databases, which were then systematically filtered using the inclusion and exclusion criteria detailed in Table 2. Figure 1 summarizes the literature search and selection procedure. The final set of 112 papers was examined to provide a thematic analysis that highlights the different applications of CDT and enabling technologies employed within each layer of a proposed conceptualized architecture.

<table>
<thead>
<tr>
<th>Table 1. Scopus search string</th>
</tr>
</thead>
<tbody>
<tr>
<td>(TITLE-ABS-KEY (&quot;digital&quot; OR &quot;virtual&quot; OR &quot;cyber&quot;) AND TITLE-ABS-KEY (&quot;twin&quot;* OR &quot;replica&quot;* OR &quot;shadow&quot; OR &quot;counterpart&quot;<em>)) AND TITLE-ABS-KEY (&quot;construction site&quot;</em> OR &quot;building site&quot; OR &quot;infrastructure site&quot; OR &quot;project site&quot;* OR &quot;construction phase&quot;* OR &quot;building phase&quot; OR &quot;construction stage&quot; OR &quot;building stage&quot; OR &quot;engineering projects&quot; OR &quot;engineering site&quot;* OR &quot;construction operation&quot;* OR &quot;construction task&quot;* OR &quot;construction activity&quot;* OR &quot;building operation&quot;* OR &quot;building task&quot;* OR &quot;building activity&quot;* OR &quot;building project&quot;* OR &quot;construction management&quot;*))</td>
</tr>
</tbody>
</table>
Table 2. Inclusion and exclusion criteria for the literature search

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Journal Articles</td>
<td>- Non-English language.</td>
</tr>
<tr>
<td>- Conference papers</td>
<td>- In press documents</td>
</tr>
<tr>
<td>- Review papers</td>
<td>- Book chapters</td>
</tr>
<tr>
<td>- Final publication stage</td>
<td>- Duplicate studies reported on research that had been previously published (e.g., a conference paper gets extended into a journal article or vice versa)</td>
</tr>
<tr>
<td>- Any publication date</td>
<td>- Studies that are irrelevant to DT concept</td>
</tr>
<tr>
<td>- English language documents</td>
<td>- Research Studies from other industries (e.g., manufacturing, aerospace, oil &amp; gas, naval, marine, aviation, urban/city planning, etc.)</td>
</tr>
<tr>
<td>- Literature covering the scope of this research (i.e., DT for construction phase applications).</td>
<td>- Studies that are out of scope (e.g., DT implementations during other phases like the design, operational, or maintenance phases).</td>
</tr>
<tr>
<td>- Literature that covers DT for the whole construction industry and includes sections related to the construction phase or construction management.</td>
<td>- Non-English language.</td>
</tr>
</tbody>
</table>

CDT Architecture and Enabling Technologies

This section introduces a conceptualized architecture of CDT to better explore and categorize the different enabling technologies and processes integrated into a CDT to serve its intended application, (see Figure 3). The proposed architecture presents five distinct layers constructing a CDT: sensing, communication, storage, analytics, and visualization layer, their data interaction, and the feedback-decision making-loop mechanism. In this context, a “layer” denotes the logical combination of the hardware and/or software components sharing common functionality and purpose. Systematically, a “CDT architecture” is an enclosed combination of the required layers constructed to interact in a way that enables data collection, transmission, processing, and visualization to support informed decision making.

Figure 3. The proposed conceptualized CDT architecture

Taxonomy of CDT Applications

Seven major CDT applications and 17 sub applications emerged from the conducted thematic analysis. This includes 1- Safety and Risk Management (SRM), 2- Progress Monitoring and Control (PMC), 3- Data Integration and Management (DIM), 4- Construction Robotics and Automation (CRA), 5- Supply Chain and Logistics (SCL), 6- Quality Control and Assurance (QCA), 7- Sustainability and Circular Construction (SCC). Figure 3 demonstrates the distribution of these major applications within the literature, providing insights into the level of research corresponding to each application. For instance, “Safety and Risk Management” emerged as the most investigated area of CDT application while “Sustainability and Circular Construction” attracted the least research attention, thus offering a promising area of CDT for future research. Table 3 summarizes the seven CDT applications, their sub-implementations, and the corresponding key technologies and processes.

Figure 2. Distribution of the identified CDT applications across the reviewed studies (total number of 90).

Figure 1. Study selection process based on PRISMA.
<table>
<thead>
<tr>
<th>Application</th>
<th>Sub-application</th>
<th>Implementations Summary</th>
<th>Key technologies and processes</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workers’ Health and Safety</td>
<td>Tracking workers’ locations and issuing proactive warnings for trespassing hazardous zones; Monitoring Worker’s physical health and posture; and Interactive site visualization for safety training.</td>
<td>RFID, LoRa, GPS, Insole pressure sensors, YOLOv4, Deep SORT, Deep learning, VR, MR headsets.</td>
<td>(Lee and Lee, 2023; Osti et al., 2021; Wu et al., 2022; Zhang et al., 2023)</td>
<td></td>
</tr>
<tr>
<td>Safety Management</td>
<td></td>
<td>Photogrammetry, kinematic sensors, spatial/orientation sensors, BIM, RFID, DNN, LST, XGBoost, BPNN, and AdaBoost.</td>
<td>(Jiang, Ding, et al., 2021; Kamari and Ham, 2022; Shariatfar et al., 2022; Zhao et al., 2022)</td>
<td></td>
</tr>
<tr>
<td>SRM</td>
<td>Identify and proactively prevent safety threats in a job site (e.g., tunnelling, crane operations, and drilling tasks).</td>
<td>Mechanical and motion sensors, inclination and load sensors, RFID, IMU, Finite Element analysis (FEM), Multi Fidelity Surrogate model, Deep learning, LoRa, Visualization interface</td>
<td>(Jiang, Ding, et al., 2022; Lai et al., 2022; Li et al., 2021; Liu, Li, et al., 2022)</td>
<td></td>
</tr>
<tr>
<td>Equipment Health Monitoring</td>
<td>Monitoring the operating status of construction equipment, mainly tower cranes, against mechanical and structural failure.</td>
<td>Bluetooth Low Energy (BLE), Real-time location system (RTLS), Bayesian network, cameras, YOLOv5.</td>
<td>(Assadzadeh et al., 2023; Huang et al., 2021; Kojima et al., 2020; Leonardo et al., 2020)</td>
<td></td>
</tr>
<tr>
<td>Human-Machinery Interaction</td>
<td>Enhancing workers-machinery interactions on site to prevent accidents and fatalities in limited space. Earthwork is the main use case</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Productivity Estimation and</td>
<td>Equipment tracking for estimating and predicting productivity and utilization rates (i.e., mainly plant equipment).</td>
<td>IMU, GPS, 3D Accelerometer, kinematic sensors, Discrete Event Simulation, ANN, cloud computing, MySQL database, Supervised learning, dashboards, CNN, RNN, Cellular Network</td>
<td>(Fischer et al., 2023; Rogaje et al., 2022; Salem and Moselleh, 2021)</td>
<td></td>
</tr>
<tr>
<td>Prediction</td>
<td></td>
<td>Onsite cameras, sound sensors, GPS, RFID, 4D BIM, YOLOv4, DNN, 3D visualization model, Web-based platform, Wireless communication, Cloud-based computing and storage</td>
<td>(Hasan and Sacks, 2021; Huang et al., 2022; Shariatfar et al., 2022)</td>
<td></td>
</tr>
<tr>
<td>PMC</td>
<td>Monitoring and most importantly visualizing the progress of certain construction activities (e.g., prefabricated element installation)</td>
<td>Laser scanning, LiDAR, Photogrammetry, Point clouds, 4D BIM, Reality capture, Object recognition, Semantic segmentation.</td>
<td>(Alizadehalechi and Yitmen, 2023; Pan and Zhang, 2021; Pour Rahimian et al., 2020; Rausch and Haas, 2021)</td>
<td></td>
</tr>
<tr>
<td>PMC</td>
<td>Capturing as-built progress (mainly point cloud data) and comparing it with as-planned data (mainly BIM data) for detecting progress deviations.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>As-Built vs As-Planned Comparison</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIM</td>
<td>Integrating DT with blockchain to enable traceable, immutable, and secure data sharing.</td>
<td>Blockchain, Smart contracts, SHA-256 algorithm, Data querying.</td>
<td>(Jiang, Liu, et al., 2021; Lee et al., 2021; Zhao et al., 2023),</td>
<td></td>
</tr>
<tr>
<td>Data Schemas-enabled DT</td>
<td></td>
<td>Ontologies, Neo4j graphs, Semantic web, IFC-schema, SPARQL.</td>
<td>(Ayinla et al., 2021; Dong et al., 2021; Schlenger et al., 2022)</td>
<td></td>
</tr>
<tr>
<td>Machinery Remote Control</td>
<td>Integrating interactive and real-time reality capturing for remotely controlling equipment, mainly plant equipment.</td>
<td>VR, MR, LiDAR, Cameras, 4G/5G networks. WebSocket, node-red UI.</td>
<td>(Hasan et al., 2022; Heikkilä et al., 2022; Hoffmann et al., 2022; Schönböck et al., 2022)</td>
<td></td>
</tr>
<tr>
<td>Robotics Tasks Management</td>
<td>Planning and supervising robotic tasks by controlling the robot’s joint angles and/or its cartesian path using a simulator platform.</td>
<td>VR, MR, Gazebo simulator, MQTT, ADS protocol, KUKA KR120, IFC-based schema, SLAM, Context-awareness, Ethernet</td>
<td>(Asadi et al., 2018; Ding et al., 2020; Liang et al., 2021; Wang et al., 2021)</td>
<td></td>
</tr>
<tr>
<td>CRA</td>
<td></td>
<td>VR, MR, Gazebo simulator, MQTT, ADS protocol, KUKA KR120, IFC-based schema, SLAM, Context-awareness, Ethernet</td>
<td>(Hasan et al., 2022; Heikkilä et al., 2022; Hoffmann et al., 2022; Schönböck et al., 2022)</td>
<td></td>
</tr>
<tr>
<td>Resource Allocation</td>
<td>Managing and optimizing the allocation of construction resources.</td>
<td>VR, MR, Gazebo simulator, MQTT, ADS protocol, KUKA KR120, IFC-based schema, SLAM, Context-awareness, Ethernet</td>
<td>(Asadi et al., 2018; Ding et al., 2020; Liang et al., 2021; Wang et al., 2021)</td>
<td></td>
</tr>
<tr>
<td>SCL</td>
<td>Integrating and visualizing logistics data including material inventory, incoming deliveries, and material consumption rates.</td>
<td>RFID, GPS, BIM data, Deep reinforcement learning (DRL)</td>
<td>(Deria et al., 2022; Gehring and Rüppel, 2023; Jiang, Li, et al., 2022; Lee and Lee, 2021)</td>
<td></td>
</tr>
<tr>
<td>Material Logistics</td>
<td></td>
<td>RFID, DRL, Dijkstra algorithm, Queuing model, Genetic algorithm, 4D BIM.</td>
<td>(Esmaeili and Simeone, 2023; Lee et al., 2022; Liu, Shi, et al., 2022)</td>
<td></td>
</tr>
<tr>
<td>Resource Allocation</td>
<td></td>
<td></td>
<td>(Esmaeili and Simeone, 2023; Lee et al., 2022; Liu, Shi, et al., 2022)</td>
<td></td>
</tr>
<tr>
<td>As-Built vs As-Designed Models</td>
<td>Ensuring the geometric quality of as-built products by verifying them with as-designed models; common with prefabricated elements.</td>
<td>Laser scanning, LiDAR, photogrammetry, MR, AR, Scan-vs-BIM, Scan-to-BIM, CNN.</td>
<td>(Rausch and Haas, 2021; To et al., 2021; Tran et al., 2021)</td>
<td></td>
</tr>
<tr>
<td>Regulatory Standards Compliance</td>
<td>Checking the quality of construction products and processes based on their compliance with relevant regulations and standards (e.g., concrete maturity properties and road compaction).</td>
<td>Piezoelectric sensor, Velocity sensor, BIM, Thermocouple sensor, FEM.</td>
<td>(Han et al., 2022; Kosse et al., 2022; Posada et al., 2022)</td>
<td></td>
</tr>
<tr>
<td>QCA</td>
<td></td>
<td></td>
<td>(Han et al., 2022; Kosse et al., 2022; Posada et al., 2022)</td>
<td></td>
</tr>
<tr>
<td>Embodied Carbon Monitoring</td>
<td>Estimating and monitoring the emission of embodied carbon dioxide during the construction process of a built asset.</td>
<td>QR code, UWB, IFC-based data schema, RFID, Accelerate sensor, BIM model, Wi-Fi.</td>
<td>(Chen et al., 2021; Mao et al., 2018; Shen et al., 2022)</td>
<td></td>
</tr>
<tr>
<td>SCC</td>
<td>Facilitating the reuse and recycling of construction materials to reduce waste and enhance the efficiency of resource utilization</td>
<td>GPS, GIS data, Monte Carlo-based simulation</td>
<td>(Züst et al., 2021)</td>
<td></td>
</tr>
</tbody>
</table>

*References:*

- Ayinla et al., 2021
- Dong et al., 2021
- Hasar et al., 2021
- Hasar and Sacks, 2021
- Hoffmann et al., 2022
- Schönböck et al., 2022
- Asadi et al., 2018
- Ding et al., 2020
- Liang et al., 2021
- Wang et al., 2021
- Deria et al., 2022
- Gehring and Rüppel, 2023
- Jiang, Li, et al., 2022
- Lee and Lee, 2021
- Lee et al., 2022
- Liu, Shi, et al., 2022
- Esmaeili and Simeone, 2023
- Lee et al., 2022
- Lee and Lee, 2023
- Osti et al., 2021
- Wu et al., 2022
- Zhang et al., 2023
- Lee and Lee, 2023
- Rausch and Haas, 2021
- To et al., 2021
- Tran et al., 2021
- (Assadzadeh et al., 2023; Huang et al., 2021; Kojima et al., 2020; Leonardo et al., 2020)
Sensing layer

The sensing layer is the first and most fundamental layer within a DT architecture that employs one or more data-sensing devices to serve as the nervous system responsible for gathering raw data from the physical environment in which they were deployed. While only devices that stream data in a timely manner should technically be considered part of the sensing layer, static data sources (e.g., BIM models, GIS, schedules, and regulations) are also included to provide a more comprehensive overview of all data sources available for supporting a DT system. The selection of data sources and sensing technologies significantly affects the accuracy, reliability, and acquisition speed of raw data, and consequently, the efficiency and overall performance of the DT system. Thus, understanding the data requirements that align with the DT's primary purpose is crucial. Figure 4a maps the different sensing technologies employed within the DT systems reviewed from the literature to the identified applications. The size of each node represents how frequently a technology was used. RFID was the most common sensing technology for safety and risk management implementations. GPS was widely used for progress monitoring, robotics, and logistics due to its real-time, precise location data for equipment, robots, and vehicles. IMUs followed GPS, mainly capturing motion data for progress monitoring and construction robotics.

Communication layer

The communication layer plays a key role in receiving the raw data acquired by the sensing layer and transmitting it to the DT hub to be stored, processed, or visualized. Typically, the communication layer utilizes either a wireless network (e.g., Wi-Fi, Bluetooth, LoRa, 5G/4G cellular, etc.), a wired network (e.g., Ethernet, USB, and Fieldbus), or occasionally a hybrid of both. Figure 4b maps the identified applications to the different employed communication networks and protocols to give insights into their suitability and popularity. For example, Ethernet is featured in most studies related to controlling scale model robotics (Asadi et al., 2018; Liang et al., 2022). By employing the MQTT protocol, Ethernet allows for bi-directional communication between the physical robot and its digital model enabling smooth robotic control (Han et al., 2022; Liang et al., 2022). Short-range wireless technologies (i.e., Bluetooth and Wi-Fi) are commonly utilized in various applications due to their higher transfer rates. Bluetooth Low Energy (BLE), a more energy-efficient choice compared to traditional Bluetooth, is also used in several studies as it is suitable for battery-powered devices and when long-period data collection with short-range coverage is required. Within long-range wireless technologies, LoRa is found to be a popular option for safety and risk management applications, specifically with hosting operations owing to its long-range coverage (~10 km) and low-power requirement personnel (Zhang et al., 2023). In general, selecting a communication network and protocols depends on several factors determined by the overall purpose of CDT including the required transmission rate, latency, range of coverage, security, and power consumption.

Storage layer

This layer provides a repository for hosting and managing the vast amount of data including the sensed raw data, data from other sources, and historical data. As shown in Figure 4c, cloud-based storage emerged as the predominant choice for data storage within CDT on account of its capacity to accommodate extensive data volumes and streamline their retrieval over the Internet (Pan and Zhang, 2021). Both SQL, relational databases, and NoSQL, non-relational databases, have been adopted in various construction applications. The selection of either relies on the types of data, its structure, and volume. For unstructured and variant data in substantial volumes (e.g., visual data), NoSQL is best due to its dynamic schema with higher scalability and adaptability. On the other hand, SQL is fit to manage well-defined data that has a structured schema (e.g., workers' information) as it gets arranged in structured tables with predefined relationships.

Analytics layer

This layer functions as the DT brain in charge of processing, analyzing, and deriving insights from the collected data to support informed decision-making. Figure 4d maps the various data analytics techniques used for each CDT application. These techniques can be streamed into three main categories: a) ML-based methods that employ deep neural networks and object recognition and tracking algorithms (e.g., CNN, RNN, YOLO, and Deep SORT). b) Numerical/Simulation-based methods (e.g., Finite element method, discrete event simulation, queuing theory, Dijkstra algorithm, and genetic algorithm) that are mostly common with applications requiring simulation and optimization (e.g., equipment health monitoring, hazards identification, resource allocation, and material logistics). c) Human interpretation which is mostly featured in quality inspection applications where defects get visually detected in as-built data (e.g., point clouds). Construction robotic applications also highlighted human supervision for planning and remotely controlling robotic tasks.

Visualization layer

This layer enables the visual representation of the processed data and insights in a dynamic and interactive environment. Different visualizations can be employed within the CDT architecture depending on the data type intended to be visualized. For example, 3D data (e.g., point clouds and BIM models) are commonly utilized to visualize progress data for quality control and progress monitoring purposes. 4D BIM models are also employed to identify and visualize hazardous activities and locations and create safety precautions. Extended reality (XR) technologies are employed in various CDT applications as they allow stakeholders to intuitively interact with visualized data in an easily accessible way. Several game engines and modelling software are featured within the literature to create 3D immersive environments. As presented in Figure 8, the “Unity engine” is the most used game engine for this purpose as it supports modeling and hosting DT entities, data integration from multiple sources, and running data analytics within its environment. Such features make it a favorable platform for visualizing DT data. Other game engines and modelling software are also highlighted, including Unreal Engine, and Blender platform. 3D visualization environments were not included in many applications (e.g., material logistics, embodied carbon estimations, and productivity measurements) as the type of data is mainly numerical and can be visualized using graphical representations within customized dashboards.
Figure 4: The different enabling technologies employed within each DT layer mapped to the identified applications: (a) Sensing, (b) Communication, (c) Storage, (d) Analytics, (e) Visualization.

SRM: Safety and Risk; Management; PMC: Progress Monitoring and Control; CRA: Construction Robotics and Automation; QCA: Quality Control and Assurance; SCL: Supply Chain and Logistics; and SCC: Sustainability and Circular Construction.
Discussion

Evaluation of CDT applications

Comparing the findings of this study with previous review studies highlights the significant progression of DT implementations during the construction phase of built projects as most prior implementations are mainly focused on the operational and maintenance stage. Additionally, most construction applications highlighted in prior review studies are mostly nD BIM applications (e.g., clash detections, cost estimation, scheduling, and communication) (Boje et al., 2020) or implementation of the integration of BIM with other technologies (e.g., XR, Simulation, AI). However, the thematic applications that emerged in this study are mainly DT-related implementations. Some of these applications align with DT research clusters categorized by (Jiang, Ma, et al., 2021) in early 2021 including safety, quality, and progress management. These applications have been expanded since then and new ones have emerged including construction robotics, logistics, and sustainability. Promising CDT research areas are recommended that are related to applications such as “circular construction”, “embodied carbon estimations”, “full-scale robotics”, and “graph-based DT”. New construction applications are anticipated to emerge, contributing to the increasing adoption of DT in both academia and industry.

What is CDT?

This section provides a discussion on the key attributes of a CDT system including the data acquisition mechanism, the feedback-decision making-loop dynamics, and the role of 3D modeling and BIM within CDT systems. Ultimately, synthesizing these discussions to propose a definition capturing the essence of CDT.

Data streaming mechanism

Almost all reviewed studies integrated at least one form of data sensing mechanism to provide DT with a dynamic stream of data without which a system cannot be considered as DT. Nonetheless, the level of automation, frequency, and latency parameters of data acquisition can all vary across different DTs based on several factors including the main purpose of the DT, the type of data that needs to be collected, and the rate at which is being updated. Several studies labelled their data-collecting mechanism as “real-time” or “near-real-time” without numerically defining these abilities in terms of latency and frequency. Only a few studies provided statistical information about these parameters to accurately reflect the capability of their data acquisition system (Deria et al., 2022; Liang et al., 2022; Posada et al., 2022). Although in many articles especially within the manufacturing industry, being able to provide real-time data with high frequency and low latency is a requirement for a system to be a DT; in many construction applications that might not be achievable or even valuable. For example, in some applications, it was adequate to collect data, mostly visual, at a frequency of once a day (Hasan and Sacks, 2023; Pour Rahimian et al., 2020; Zhao et al., 2022). Yet, in other applications where data gets updated more dynamically, a higher frequency is essential. For instance, for tracking the angle movement of a robotic arm, a frequency of up to 250 Hz might be required to synchronize with the rapid rate of update of the moving arm (Liang et al., 2022).

In general, acquiring vast amounts of construction data with high frequency might not be needed and comes with high costs and significant environmental impact associated with collecting, transmitting, storing, and processing this data. Therefore, when designing a DT, these attributes should be optimized to create a balance of being efficient and meeting its overall purpose requirements. Similarly, the degree of autonomy in data acquisition varies widely, with some cases featuring semi-automated approaches, particularly in laser scanning and photogrammetry where human intervention is required for capturing and processing the visual data. In contrast, many studies employ fully automated sensing sources, like positioning sensors capable of transmitting data without any human involvement.

This is to emphasize that the spectrum of data acquisition capabilities for DTs is broad, and achieving a fully automated data stream within specific latency and frequency parameters is not always a must-met criterion for classifying a system as a DT.

Feedback loop—decision making—dynamics

The majority of the DT research in digitally advanced fields (e.g., manufacturing) emphasizes automated bi-directional communication between the physical entity and its digital counterpart (Ladj et al., 2021). In manufacturing, its typically well-designed and controlled environments (e.g., an assembly line) enable the establishment of an automated feedback loop. However, with dynamic, variant, and unpredictable construction operations, establishing an automated feedback channel from the DT to the site can be challenging. Instead, the feedback in some construction applications often involves visualized insights presented to stakeholders for decision-making, highlighting a human-oriented feedback channel, (see Figure 3). For instance, within progress monitoring applications, site stakeholders are presented with performance metrics, which they then align with baseline targets to inform their decision-making.

Almost all construction robotics and automation applications feature fully automated feedback channels where control instructions are given back to a prototype machinery or a robot to conduct a desired task based on the acquired data. Nonetheless, most of these studies used scaled machinery or robotic assembly to validate their systems, raising concerns about their scalability and real-world construction site applicability. Automated feedback was also employed in some safety applications in the form of proactive visual or audible warnings automatically triggered when a potential hazard is detected (Jiang, Ding, et al., 2022; Zhang et al., 2023).

BIM-CDT relationship

The relationship between DT and BIM has been a subject of debate among researchers, with some viewing DT as an extension of BIM with the incorporation of new technologies (e.g., IoT, AI, and simulation), while others perceive them as two distinct concepts with significant differences in their characteristics and applications. In many applications where objects’ non-geometric information (e.g., attributes and relationships to other objects) are the main data of interest rather than the geometric information, BIM may not always be necessary for DT functionality and simple 3D models with low Level of Detail (LoD) would be enough for visual representation.
Examples of such applications are workers’ health monitoring where workers’ information and status are of interest or in material logistics where material quantities or locations are tracked. Nevertheless, in several applications, BIM models were still integrated into CDT layers, serving different purposes as follows: a) within the sensing layer as a source of static data (e.g., schedule, quantities, semantic and geometric data); b) within a storage layer as a unified storage model that integrates and stores various relevant data; c) provides a 3D visualization and interactive environment, possibly integrated with dashboards, within the visualization layer.

A proposed CDT definition

Several definitions of DT exist within the construction industry, often borrowed from advanced sectors like aerospace and manufacturing, where the term originated and developed. These definitions typically include terms like virtual replica, 3D model, digital counterpart, realistic model, and real-time representation of physical. However, a critical question arises: Is DT primarily about high LOD modeling and replicating physical objects, or is it fundamentally about data management and informed decision-making? Particularly in construction implementations, the focus is more on data capturing and analysis that enables informed decision making rather than 3D replication or modeling. Then, there is a need for a definition that emphasizes the data management aspect of DT in construction. Drawing insights from the extensive literature review and the conceptualized DT architecture in Figure 3, this study proposes the following definition:

Construction Digital Twin (CDT) is a system responsible for collecting and processing construction dynamic data about a physical entity within the intended application’s temporal demand, translating this data into actionable knowledge to enable informed decision making.

“Dynamic data” refers to data that is continuously changing and collected through automated sensing mechanism. It is important to note that CDT can also combine static data from BIM models or historical records to supplement the dynamic data when needed. A “physical entity” can be a tangible object (e.g., a structure, equipment, or person) or a physical process (e.g., material delivery or element installation). “Temporal demand” acknowledges the varied data latency and frequency requirements across different DT construction applications, as discussed earlier. Consequently, terms like “real-time” and “near real-time” are avoided to highlight the variability in data capturing and processing rates and the lack of precise numerical definitions for these terms. Informed decision-making can be either fully automated or human-assisted, as illustrated in Figure 3.

Literature gaps and recommendations for future research

- Many of the reviewed studies relied on a single sensing source for data acquisition mainly to avoid the challenges of integrating data streaming from multiple sources with different formats. Still, it is frequently stated that depending on a single source for onsite data falls short of providing accurate and comprehensive insights into scalable operations (Hasan and Sacks, 2023; Hoffmann et al., 2022; Pan and Zhang, 2021; Salem and Moselhi, 2021). Only by merging data from multiple sources, a better understanding of construction operations and better-informed decision-making is obtained. Hence, it is recommended to aim at integrating heterogeneous data across multiple domains, leveraging sensor fusion to bridge this gap. Moreover, publishing well-defined data schemas for construction applications can facilitate representing complex information and potentially contribute to developing industry standards.

- Another gap is the absence of standardisation in DT design and deployment including the configuration and specifications of key parameters including data acquisition frequency and latency, analytics accuracy, and its degree of autonomy. Most studies loosely labelled their DTs as “real-time” or “near-real-time” systems without numerically defining those parameters and linking their defined attributes to the DT's main purpose.

- Besides, many studies showcased a limited scope by carrying out a single data collection trial, failing to illustrate how the data and analyzed metrics are regularly updated in a timely manner. This raises questions about the usability, scalability, and versatility of their DT systems. Standardizing DT implementation is key for facilitating its adoption in the industry as it ensures consistency and compatibility.

- Several studies relied on human interpretations for analysing the collected data. This is especially the case with most of the quality inspection and progress tracking applications. Even when automated data analytics are integrated, they are limited to presenting descriptive analytics (“What has happened? As-is analysis”). Only a few studies employed predictive analytics (Lee and Lee, 2021; Li et al., 2021). It is recommended for future research to gear towards enabling CDT to provide predictive (“What will happen? -To-be analysis”) and prescriptive analytics (“What should be done? -What-if analysis”) by integrating advanced AI and simulation techniques. Doing so will unlock the complete potential of CDT as a proactive decision-support system.

Conclusion

This study serves as a valuable resource for researchers and practitioners in the construction industry by presenting a structured taxonomy of CDT applications and their interaction with different technologies through systematic review and thematic analysis. Seven main thematic applications of CDT are identified and discussed. Additionally, a conceptualized architecture of CDT, comprising five distinct layers, is proposed to aid in exploring and analyzing enabling technologies and processes, and mapping them to the identified applications. The paper also explores and discusses the characteristics that define the essence of CDT and proposes a definition emphasizing its focus on data management and decision-making. These findings not only consolidate existing knowledge in the CDT domain but also provide a foundation for guiding future research and development in this field.

References


