# What is a Digital Twin anyway? Deriving the definition for the built environment from over 15,000 scientific publications

Mahmoud Abdelrahman [a], Edgardo Macatulad [a,b], Binyu Lei [a], Matias Quintana [c], Clayton Miller [d], Filip Biljecki [a,e],*

[a] Department of Architecture, National University of Singapore, 4 Architecture Drive, Singapore, 117566, Singapore
[b] Department of Geodetic Engineering, University of the Philippines, Diliman, Quezon City, 1101, Philippines
[c] Future Cities Lab Global Programme, Singapore-ETH Centre, CREATE campus, #06-01 CREATE Tower, Singapore, 138602, Singapore
[d] Department of the Built Environment, National University of Singapore, 4 Architecture Drive, Singapore, 117566, Singapore
[e] Department of Real Estate, National University of Singapore, 15 Kent Ridge Dr, Singapore, 119245, Singapore

## ARTICLE INFO

## ABSTRACT

The concept of Digital Twins (DT) has attracted significant attention across various domains, particularly within the built environment. However, there is a sheer volume of definitions and the terminological consensus remains out of reach. The lack of a universally accepted definition leads to ambiguities in their conceptualization and implementation, and may cause miscommunication for both researchers and practitioners.

We employed Natural Language Processing (NLP) techniques to systematically extract and analyze definitions of DTs from a corpus of more than 15,000 full-text articles spanning diverse disciplines. The study compares these findings with insights from an expert survey that included 52 experts. The study identifies concurrence on the components that comprise a "Digital Twin" from a practical perspective across various domains, contrasting them with those that do not, to identify deviations. We investigate the evolution of digital twin definitions over time and across different scales, including manufacturing, building, and urban/geospatial perspectives. We extracted the main components of Digital Twins using Text Frequency Analysis and N-gram analysis. Subsequently, we identified components that appeared in the literature and conducted a Chi-square test to assess the significance of each component in different domains.

Our analysis identified key components of digital twins and revealed significant variations in definitions based on application domains, such as manufacturing, building, and urban contexts. The analysis of DT components reveal two major groups of DT types: High-Performance Real-Time (HPRT) DTs, and Long-Term Decision Support (LTDS) DTs. Contrary to common assumptions, we found that components such as simulation, AI/ML, real-time capabilities, and bi-directional data flow are not yet fully mature in the digital twins of the built environment. We derived two definitions for the Building/Architecture DT and the City/Urban DTs. Both definitions have a must-have components (such as spatial and temporal data updates) and good-to-have components such as prediction, AI, bi-directional data flow, and Real-time data exchange.

One of the key findings is that the definition of digital twins has not yet reached its equilibrium phase, highlighting the need for ongoing revisions as technologies emerge or existing ones become obsolete. To address this, we introduce a novel, reproducible methodology that enables researchers to refine and adapt the current definitions in response to technological advancements or deprecations.

---

* Corresponding author at: Department of Architecture, National University of Singapore, 4 Architecture Drive, Singapore, 117566, Singapore.
E-mail address: filip@nus.edu.sg (F. Biljecki).

## 1. Introduction

The concept of a Digital Twin (DT) traces its origins to a book titled Mirrored Worlds published three decades ago [1]. Various terms have been coined subsequently, e.g. Mirrored Space Model (MSM) and Information Mirroring Model (IMM), and later "Conceptual Ideal for Product Lifecycle Management (PLM)" by [2], evolving the discourse with defining the main components of a DT. Since then, the term has evolved to encompass a broad spectrum of concepts and levels of fidelity, ranging from static models to dynamic simulations with real-time data integration and even autonomous capabilities (from a data visualization tool to a fully autonomous decision making system such as autonomous cars). For example, the term and concept of Digital Twin have been widely adopted in manufacturing industry [3,4], and transcended it, reaching the built environment and urban/geospatial domains [5–17] including urban logistics and transportation [18, 19], urban–social interaction [20], participatory human-centric exploration [21,22], conversion of Building Information Modeling (BIM), IoT, and GIS [23,24]. Furthermore, according to more recent literature, the DT scope appears to have broadened to represent the entire lifecycle of a system, not just a single lifecycle phase (e.g., conceptual, design and planning, construction/manufacturing, operation and maintenance, and decomposition) [25] as well as a wide range of applications [26]. At the time of writing this paper (mid-2024), a query for the term 'Digital Twin' on ScienceDirect, a major database of peer-reviewed scholarly literature, yields 52,495 results. The majority of these results are dominated by Smart Manufacturing and Industrial Engineering (e.g., Industry 4.0). On a separate query, a total of 11,000 results are found to be related to the Built Environment (BE)[1].[2] This multitude of literature demonstrates the intense and rapid interest in digital twin concept and technology and its applications in the built environment [24,27–29]. Despite the growing interest in DTs within the built environment, the concept has been reported to remain unclear for many practitioners and researchers [30,31], unlike in some other industries such as Manufacturing [32] and Aerospace [33] where it appears to be converging toward a well-established consensus.

The rapid growth of interest in emergent technologies such as DTs can lead to fragmented knowledge about its core functionalities and key elements. This fragmentation often results in ambiguities across industry and research papers concerning its precise definition [32,34–37], its essential components and minimum functionalities for a system to be considered a DT, and how it is applied specifically within the BE. For example, in discussions surrounding DTs, terms such as 'data reference', 'software representation', 'living model', 'virtual replica', and 'visual asset management system', '3D city models' and 'City Information Models (CIM)' are all used interchangeably. However, not all of them accurately capture the exact essence of a Digital Twin. Amplifying this confusion, researchers have defined DTs from one or more perspectives, often vaguely, leading to a range of interpretations ( Table 1). Further, the ambiguity is exacerbated by its common use as a buzzword and the term is also affected by the expedient practice and trend of giving new names to existing concepts, simply to make them seem novel.

On the other hand, many researchers have defined the lowest common denominator of the Digital Twin concept [38–41], which consist of 3 main components: (1) Physical entity, (2) Virtual entity, and (3) Data connection. Others have extended these minimum components to

include (4) Real-time capabilities, (5) Real-time update of the digital system with the physical system, and (6) Decision support [42–44] (Fig. 1). While these interpretations may not always provide a comprehensive understanding, and some may even collide, they can capture various aspects of digital twins relevant to the specific study where the definition is employed. By consolidating diverse perspectives on digital twins from a large number of sources, our goal is to formulate a data-driven, comprehensive, and standardized definition of digital twins that captures a unified understanding across the built environment, and bring us one step closer to a consensus.

Considering the inconsistent terminology and contrasting implementations, recent studies have introduced the concept of DT maturity stages [45–48], which is used to assess how advanced a DT is in terms of technology and functionality. These may range from basic and static geometric models to highly integrated systems with advanced features such as real-time capabilities and AI/ML functionalities supporting simulations. For example, [45] and [49] identified 5 maturity stages in City DTs: (1) Reality capture, (2) 2D maps or 3D models, (3) addition of real-time data, (4) two-way integration and interactions, and (5) autonomous operations and maintenance. [46] introduced the CITYSTEPS Maturity model which grades them according to eight maturity stages: (1) Preparation and planning stage; (2) 2D stage; (3) 3D static stage; (4) 3D dynamic stage; (5) Dynamically Integrated 3D Stage; (6) Real-time decision-making stage; (7) Autonomous decision-making stage; and 8) Real-time synchronization and autonomous implementation stage. [50] conducted a review on CDTs, highlighting their maturity levels and differentiating them from traditional 3D city models [51–53], while identifying key research gaps, such as the need for advanced real-time data analytics and public participation in CDT development. These recent developments signal that the community is increasingly acknowledging that DTs may come in different shapes and flavors and considering Digital Twin as a broader umbrella term for a range of concepts and implementations of virtual replicas of the built environment. Nevertheless, an overarching definition and conceptualization of the term in the built environment remain missing [38,54,55] despite the substantial efforts invested to address this gap [12,56–59].

In this paper, we conduct a large scale systematic review coupled with an expert survey to shed light on the definition of a digital twin in the built environment. A systematic review methodology was used because it helps reduce bias in the selection and analysis of existing research, and it follows initiatives in other fields such as circular economy that have also grappled with similar terminological issues [60,61]. This review involves analyzing approximately one thousand Digital Twin mentions, interpretations, explanations, and definitions that have appeared in more than 15,000 publications from 2001 to January 2024. By systematically examining such a large sample size, we provided an empirical snapshot of current scholarly understandings of DTs, as well as the evolution of the definition over the years. We construct our approach on the premise that researchers are likely to choose definitions that align with their research focus or objectives, even if some of them do not align with the current predominant definitions. We then analyzed the definitions used in the Built Environment (Building, Urban domains), and we used Manufacturing domain as a benchmark for comparison to extract the DT components of each. The analysis aimed to understand the various perspectives on DTs, develop a definition based on the most common characteristics understood by researchers through practice and application, and ensure this definition was comprehensive enough to encompass a wide range of DT components, ultimately capturing the true essence of what constitutes a Digital Twin or in other words, of what BE researchers have been referring to as Digital Twins. We also consider the maturity levels of DT in different domains, recognizing that not all implementations require the same level of maturity. Finally, by deriving a definition, in our paper we question some instances in practice that manipulate the definition of DT to align their concepts as DT but significantly depart from the actual concept.

---

[1] We obtained this estimation by using the following boolean search: *"('Digital twin' AND ('Building' OR 'Built Environment' OR 'City' OR 'urban' OR 'smart building' OR 'BIM' OR 'architecture')"*

[2] In the context of this paper, we refer to the BE as buildings, urban spaces/places and cities. which constitute different scales within the Architecture, Engineering, and Construction (AEC) domain
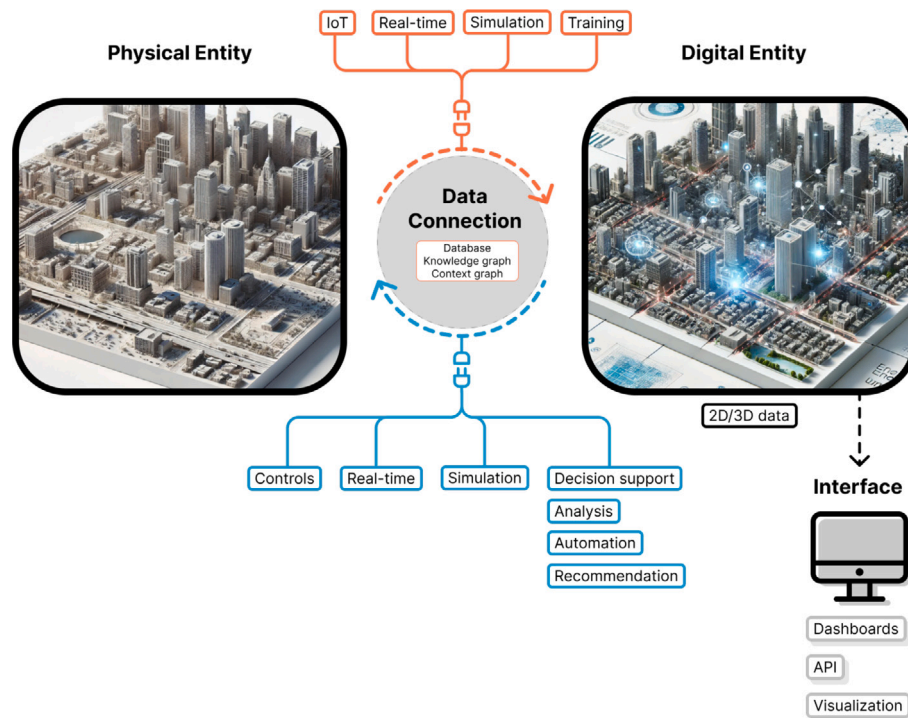
**Fig. 1.** Illustration of commonly discussed Digital Twin components.

The approach to obtain clarity on terminology from a vast body of knowledge that we introduce in this paper is novel and our development, and a contribution that can be applied also in other domains. The rest of the article is structured as follows. Section 2 provides a review of the literature on Digital Twins, highlighting the various definitions that have been proposed in the literature. Section 3 outlines the methodology used in this research, including data sources, data processing, and the methodology to derive a new definition of Digital Twin. Section 4 presents the results of the study, including the analysis of the components and definitions extracted from the literature and the expert survey. Section 5 discusses the findings, implications, and futures of the study, and Section 6 concludes the paper.

## 2. Related work

Numerous surveys and documents on DTs have been undertaken across various disciplines, with a notable emphasis on manufacturing and industrial sectors. In the context of manufacturing, ISO 23247-1:2021 defined DT as "fit for purpose digital representation of an observable manufacturing element with a means to enable convergence between the element and its digital representation at an appropriate rate of synchronization" [62]. [63] at NASA defined DT as "an integrated multi-physics, multi-scale, probabilistic simulation of an as-built vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its corresponding flying twin".

Table 1 presents a list of review studies conducted on DTs and the number of definitions of DTs in each study. Some of these surveys do not primarily focus on defining DTs but rather their history and sometimes mention one or more definitions that are widely recognized or commonly used. Some papers have, however, conducted surveys that, for a large part, focus on characterizing and defining the DT concept [27,64–67], but these are not focused on the BE domains.

The different implementations across domains also reflect the varied definitions of DTs. Different DT architectures or frameworks for different use cases conceptualize and specify the respective DT components to be implemented. [39] introduced the DT concept model in

manufacturing that is composed of 3 main parts: (1) physical products (i.e., in real space), (2) virtual products (i.e., in virtual space), and (3) connections that link the virtual and real. [95] expanded this concept model into a five-dimensional DT model, presented to be a reference to understand and implement the enabling technologies and tools for DTs. Their model is composed of (1) physical entities — which may be any device, product, physical system, or process; (2) virtual models — which are the replicas of physical entities (including their physical geometries, properties, behaviors, and rules); (3) DT data — referring to the heterogeneous data that drives the DT (static attributes or and dynamic data obtained from physical entities, as well as data generated by virtual models e.g., from simulation); (4) services, which capture the paradigm of Everything-as-a-Service (XaaS), directing the applications and use case of the DTs; and (5) connections, which specifies the links that enable the other four parts to collaborate.

In urban applications, DT is seen as one of the enabling factors for digital services (i.e., smart cities) and for modeling of urban processes [96,97]. Such DT implementations enable simulations that can be used for analysis and prediction through scenario development [98]. This technical aspect of digital twin of cities is oftentimes the scope of DT review [30,95,99,100], however, this perspective sometimes lead to its biased understanding [101]. The Digital Urban European Twins (DUET) project[3] is an example of an established DT initiative for cities that promotes the adoption of local and urban DTs. The technical report on the architecture of the DUET system [102] presented the overall architecture of their platform and the specifications of the DUET system in terms of the functionalities of the respective processes and the corresponding components that realize them. The DUET architecture refers to the system's components as multiple distinct layers: (1) presentation layer — contains all user interfaces and visualization modules; (2) access control layer — contains all necessary *Gateway* components that control data flow; (3) service layer — exposes the functionalities of core components through a set of APIs; (4) business layer — tackles user

---

**Table 1**
Digital Twin related literature review papers and the number of definitions in each.

| Category | Ref | No. of definitions |
|---|---|---|
| Manufacturing | [38] | 5 |
| | [64] | 30 |
| | [68] | 6 |
| | [65] | 16 |
| | [34] | 6 |
| | [69] | 1 |
| | [70] | 1 |
| | [71] | 3 |
| | [72] | 1 |
| | [35] | 1 |
| Smart Cities, Construction, and Built Environment | [27] | 25 |
| | [73] | 4 |
| | [31] | 1 |
| | [74] | 8 |
| | [30] | 4 |
| | [75] | 1 |
| | [76] | 1 |
| | [77] | 1 |
| | [14] | 1 |
| | [56] | 1 |
| | [78] | 3 |
| Energy | [79] | 4 |
| | [80] | 1 |
| | [81] | 1 |
| Transportation | [82] | 7 |
| | [83] | 10 |
| | [84] | 1 |
| Healthcare | [85] | 6 |
| Other Industries | [36] | 7 |
| | [86] | 1 |
| | [87] | 1 |
| General Topics | [88] | 1 |
| | [66] | 47 |
| | [89] | 5 |
| | [90] | 1 |
| | [91] | 3 |
| | [92] | 1 |
| | [93] | 3 |
| | [94] | 4 |

needs; (5) data layer — consists of all the heterogeneous data repositories, including information about data sources, ontologies, users, and other metadata; (6) infrastructure layer — consists of physical or virtual servers, and other relevant supporting tools; and, (7) security layer — implements necessary security measures for all layers.

The concept of DTs continues to gain attention across different fields, particularly for its ability to continuously analyze incoming (i.e., real-time) data and perform data-driven simulations [26,103]. As different domains recognize the potential of DTs and move toward implementations for various applications, standards play an important role in the creation of DTs, especially concerning data from different sources for use in various applications [58]. For complex systems such as cities, the practical approach of digital twinning would be to begin with priority areas and scale over time which would require common frameworks for effective communication [97].

Defining standards – which may relate to one or several elements of a DT and of its construction – is the scope of ongoing work [91]. Recently, the joint technical committee of the International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) published the standard *ISO/IEC 30173:2023, Digital twin — Concepts and terminology* which presents general descriptions of DT concepts, terminology and application overviews and aims to provide a common basis for understanding its concept and composition and be applicable in any and across domains [104]. In the standard, a DT system is described as a *hybrid entity* or *system of systems*, with the DT elements illustrated within this context, which include the Target

entity, which refers to something (physical or non-physical) having a distinct existence and is the subject of digital representation; the Digital entity, which is the computational entity comprised of data and procedural elements; and the Data connection between the two, and corresponding models, data and interfaces (e.g., APIs) involved— mirroring the DT concept model of [39]. Beyond these three, the standard further describes the other elements of the DT system as the Services, which refer to the capabilities (e.g. analysis, simulation, visualization, optimization, etc.) of the DT, the Infrastructure, which include the supporting components to ensure the DT's operation (e.g., sensors, network, GIS, etc.) and the System aspects, which are concerned with the integration, governance, security, and privacy, and data ethics involved in the DT implementation. However, there has yet to be a DT reference model that unifies the *ingredients* necessary to build a DT [97]. Thus, alongside unifying definitions, it is equally important to identify and unify the enabling components of DT development and implementation, especially toward goals of interoperability and scalability.

While several academic reviews have attempted to identify or define DTs, these definitions are often aligned with specific research domains or adapted to fit particular applications. This multiplicity has resulted in a variety of definitions, some of which accurately represent the concept of a DT, while others do not and diverge from the mainstream. There is no single definition that fits all domains, and most existing definitions are meant to be universally applied. However, our research aims to redefine DTs on the basis of the researchers' interpretation of the term within each specific domain, and we also challenge existing instances of (ostensible) DTs as we noticed papers that attempt to forcibly fit certain instances into the DT concept. In doing so, we evaluate and establish definitions for each domain individually based on the essential components unique to each field and its users.

## 3. Methodology

A coarse level of agreement exists on what qualifies as a Digital Twin, though the threshold between what a DT is or is not can be quite nebulous across different applications. To address this issue, we aim to accumulate as many insights as possible about interpretations of DTs from the current state-of-the-art. We then define the boundary between what is considered a DT and what is not, driven by data and statistical analysis.

Our approach leverages the credibility of existing literature, which has often undergone double-blind reviews by experts in various domains. This lends validity to several studies that qualify as a Digital Twin, even if they do not match standardized definitions. Accordingly, our methodology is grounded in both empirical observations and the extensive body of existing literature, combining theoretical insights with practical validation (Fig. 2). To validate our outcome, we conducted an expert survey from a panel of 52 experts in the BE DT domain. Additionally, we considered different maturity stages of DTs as defined by [45,46].

In this section, we first present our approach to identifying the sample from various resources, including full-text publications and the expert survey. Then, we outline the methods used to extract and process approximately 800 definitions and interpretations of DTs from this sample. Our procedure includes a variety of NLP methods such as Regular Expressions (Regex) [105], Sentence Embeddings [106], Frequency Analysis and N-gram analysis [107], and Fuzzy Matching [108]. Additionally, we employ Machine Learning approaches to group and cluster semantically similar definitions using *k*-means clustering and Large Language Models (LLMs). After that, we describe the methods used in extracting the DT components from different scales within the Architecture, Engineering, and Construction domain; namely, Building, Architecture, and Urban scales; in addition to Manufacturing domain as a dominant domain in this area. We used word cloud analysis and LLMs for that purpose. Then we test the significance of each component using
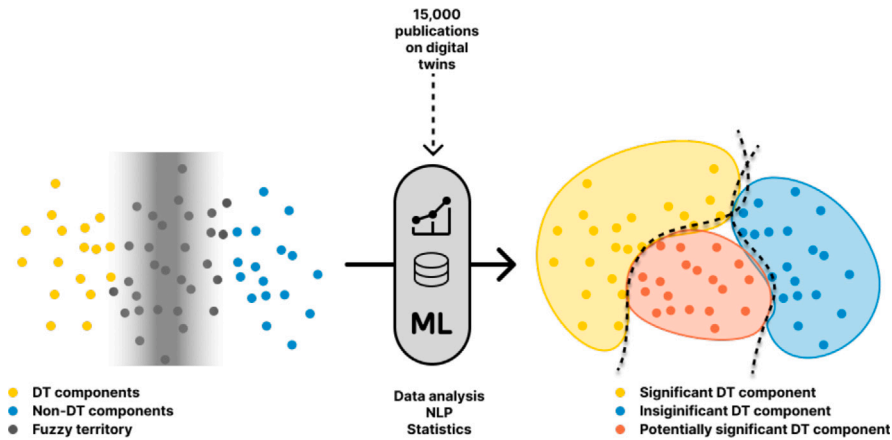
**Fig. 2.** Our novel approach to infer a DT definition consists of deriving significance of each DT component from literature using Natural Language Processing (NLP) and statistical analysis.
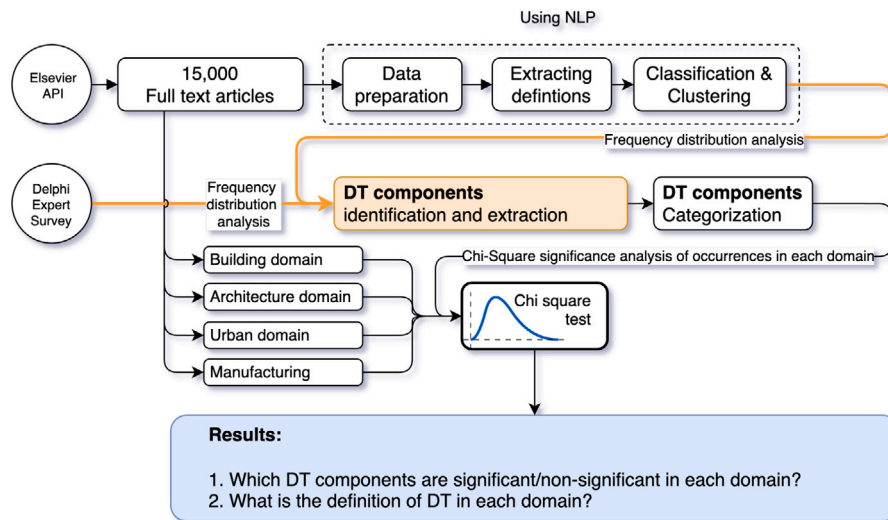


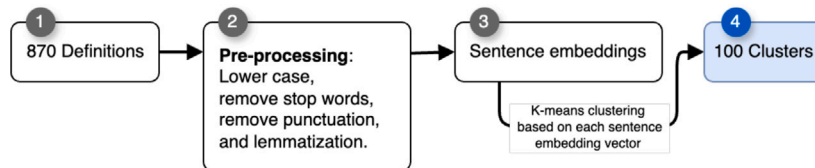**Fig. 3.** Overview of the methodology.



**Fig. 4.** The flowchart represents the definition of extraction and clustering methods.

Chi-square analysis. We maintained our methodology as systematic as possible to ensure it is generalizable to other research domains, which may also benefit from standardizing and analyzing terminology. Thus, the pipeline and code have been openly released on GitHub.[4] Figs. 3 and 4 summarize the methodology used in the current study.

---

### 3.1. Sample development

This study is built on two main data sources. The first consists of (1) full-text articles, while the second involves a (2) expert survey which is used to validate the results from a Built Environment expert perspective. From these two datasets, we have derived: the (3) DT definitions dataset and the (4) DT components dataset. In this section, we overview our methodology of collecting the sample data and extracting
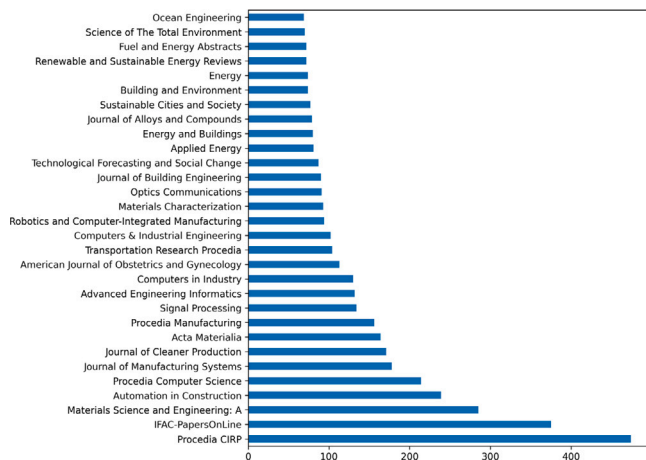
**Fig. 5.** Most common publication sources in the dataset, spanning multiple communities, scales, and domains.



**Fig. 6.** Distribution of subjects in the selected dataset. The total search returned 15,353 results; however, this figure represents a subset of articles with valid subject metadata relevant to the analysis after cleaning the data.

the definitions' and components' datasets that are used for analysis. We keep our methodology as detailed as possible for the purpose of reproducibility and reusability across different domains and studies.

### 3.1.1. Full-text articles dataset

Full-text samples were acquired using an Application Programming Interface (API) provided by Elsevier,[5] which provides interfaces to Scopus, Science Direct, and other Elsevier products. We chose to focus on the Elsevier API because it is one of the largest publishers of scientific literature and provides a comprehensive database of peer-reviewed literature across myriads of fields, thus well capturing the body of knowledge without bias. We used the API to search for articles containing the following terms: "Digital Twin", "City Digital Twin", "Urban Digital Twin", "Smart City", and "Building Digital Twin" in the title, abstract, or keywords. The search was conducted in English and covered articles published between 2000 and 2024. The search returned a total of 15,353(Fig. 6). The dataset was also collected from various sources (Fig. 5), notably the top are from manufacturing (e.g, IFAC Journal of Systems and Control, Journal of Manufacturing Systems, Procedia Manufacturing); and the Built Environment (e.g, Automation in Construction, Transportation Research Procedia, Journal of Building Engineering, Energy and Buildings, and Building and Environment).

### 3.1.2. The expert survey dataset

We compare the results of systematic review to a question — 'What is a digital twin?', collected in a Delphi survey distributed to experts. The definitions of DTs vary from different fields, making it challenging to achieve consensus in the built environment. Therefore, we built on top of a previously conducted Delphi survey by [26] to gather insights from practitioners actively engaged in this line of research. In this work, the collected responses of DT definitions are used to understand different perspectives in practice, from industries to governments. We incorporate the responses as a validation dataset, aiming to evaluate and compare with the developed corpus of research articles, thereby providing insights into the understandings of DTs. We deem that the results of this specific question can contribute to a practical perspective to supplement the analysis of the literature review and, meanwhile, detail differences between key components and focus.

The expert survey invited a structured panel of experts, who have domain expertise in DT related to the Built Environment, aiming to
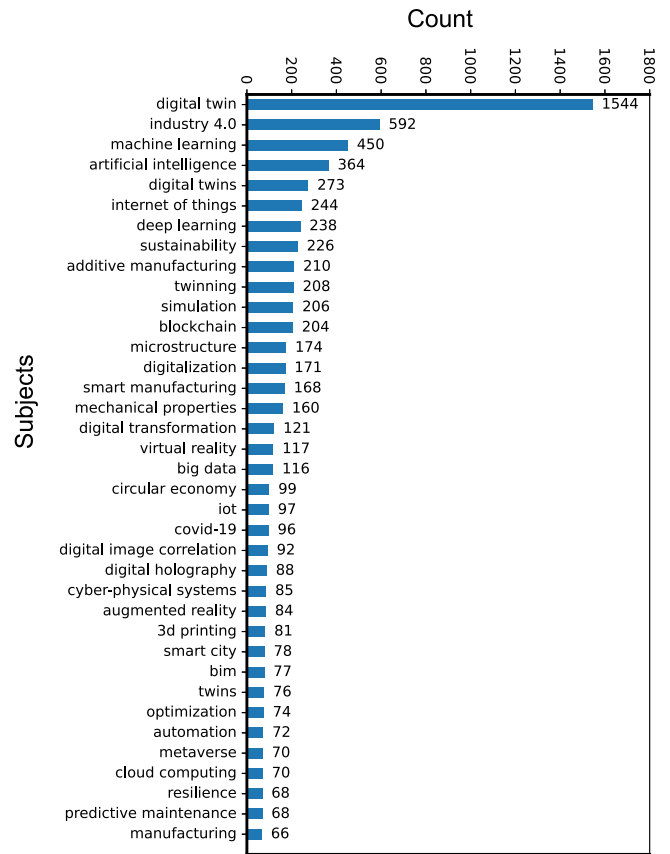
generate a consensus on the definition of DTs. A total of 52 international experts from different sectors, such as government, industry (e.g. companies, consultancies), and non-government organizations (NGOs), were selected to build the panel, contributing with diverse backgrounds and experiences in this topic. In particular, the panel of domain experts represents 23 countries, with the majority working in Europe. Of these participants, 14 were from industry, 12 were from government, and the remaining 23 were affiliated with universities or research institutes. Notably, 44% of respondents reported over 20 years of experience in fields integral to digital twins. Through an online questionnaire, all participants were asked to identify a digital twin based on their understanding and practice.

### 3.1.3. The definitions dataset

We used Regular Expressions (regex) to extract the mentions of DTs from the dataset, based on the reasoning that a subset of instances will include definitions. Regex has the advantage of extracting patterns of sentences instead of searching for an exact word. For example, among other regex patterns, we used the following pattern to extract definitions:

```
r'\b(?:digital twin(?:s)?)\s+(?:is\s+)?(?:
    defined\s+as)\s+.+?(?=\.\s|\n|$)'
```

This pattern extracts all sentences that match "digital twin/s (is/are/-can be/could be) (defined as/described as/ characterized by) …". Many of the extracted definitions are valid, and some are not valid definitions such as: "Digital Twin is a revolutionizing technology in the industry 4.0 era" and "Digital Twin is defined as a new paradigm in simulation". Therefore, we applied NLP and LLMs to filter the dataset, aiming to address two main issues: *Incomplete* and *Fuzzy Duplicated* definitions.

---

**Table 2**

Ten examples of extracted DT definitions. The rest of definitions can be found on the GitHub repository.

| Definition | Ref. |
| --- | --- |
| Digital Twin is defined as an integrated multi-physics, multi-scale, probabilistic simulation of a vehicle or system that uses the best available physical models, sensor updates, fleet history, and so forth, to mirror the life of its physical twin. | [115] |
| Digital Twin is defined as a software representation of a physical asset, system or process designed to detect, prevent, predict and optimize through real time analytics to deliver business value. | [116] |
| Digital Twin is a highly dynamic concept growing in complexity along the product life cycle, which is more than a pool of digital artifacts; Digital Twin has a structure consisting of connected elements and meta-information as well as semantics. | [117] |
| Digital Twin is a set of virtual information constructs that fully describes a potential or actual physical manufactured product from the micro atomic level to the macro geometrical level. | [118] |
| Digital Twin is a collection of all digital artifacts that accumulate during product development linked with all data that is generated during product use. | [119] |
| Digital twin is a living model of the physical asset or system, which continually adapts to operational changes based on the collected online data and information and can forecast the future of the corresponding physical counterpart. | [120] |
| Digital twin is a digital representation of a physical asset reproducing its data model, behavior, and communication with other physical assets. | [121] |
| Digital twin is a detailed simulation model of a system, for which application-specific data is stored to describe the real twin. | [122] |
| Digital twin is a dynamic representation of a physical system using interconnected data, models, and processes to enable access to knowledge of past, present, and future states to manage action on that system. | [123] |
| Digital twin is a real-time, virtual replica of a physical object or system created using sensors and computational models. | [124] |

To improve the quality of the dataset, the filtering process was designed to address these two issues. *Incomplete Definitions (ID)* refer to descriptions, explanations, or specifications that lack the necessary or essential details to fully define or explain the DT concept [109]. Examples include: "Digital twin is a revolution in the digital industry" [110], "digital twin is defined as a virtual replica of a real system" [111], "digital twin is defined as a digital representation of assets, processes, or systems" [112], and "digital twin is a living model that drives a business outcome" [113].

*Fuzzy Duplicated Definitions (FDD)* refer to instances where there are multiple similar or nearly identical definitions or descriptions, but with slight variations or differences [114].

To perform the filter, we first excluded all the incomplete definitions using LLMs, namely, the Generative Pre-trained Transformer (GPT4) model. Then, we manually validated a subset of the results.

Secondly, to overcome the FDD issue, we first clustered the definitions using sentence embeddings [125] and *k*-Means clustering into three iterations of clustering 400, 100, and 50 clusters. The Sentence Embeddings first convert all the definitions into a lower dimensional vector where the position of each definition in the vector space represents its semantic meaning. By doing this, the closer the definitions in the vector space, the more semantically similar these sentences are to each other. Each sentence is represented as a feature vector $\bar{X}_d = [x_0, x_1, x_2, \ldots, x_n]$ where $X_d$ is the definition sentence vector. We used the cosine similarity distance to measure the similarity between different definitions in the definition dataset. The cosine similarity distance between two vectors **X** and **Y** is given by:

$$\text{cosine\_similarity}(\mathbf{X_{d1}}, \mathbf{X_{d2}}) = \frac{\mathbf{X_{d1}} \cdot \mathbf{X_{d2}}}{\|\mathbf{X_{d1}}\| \cdot \|\mathbf{X_{d2}}\|} \tag{1}$$

where:

- · represents the dot product of two vectors.
- $\|\mathbf{X_{d1}}\|$ and $\|\mathbf{X_{d2}}\|$ represent the Euclidean norms (lengths) of the vectors $\mathbf{X_{d1}}$ and $\mathbf{X_{d2}}$ respectively.

After clustering the definitions into groups, we applied the Fuzzy Matching algorithm within each group to figure out how each pair of definitions matches. By doing so, we excluded semi-duplicated definitions. For example, the following two definitions are 92% matching while the difference is only in the words in italics ("virtual representation" and "digital replica"):

- "Digital Twin is a *virtual representation* of real-world entities and processes, synchronized at a specified frequency and fidelity" which was introduced by in the Digital Twin Consortium in 2020 [126].
- "Digital Twin is a *digital replica* of real-world entities and processes, synchronized at a specified frequency and fidelity", which was found in the Ref. [127].

The fuzzy matching algorithm between two definitions can be represented as follows:

$$\text{Fuzzy\_Match}(s_1, s_2) = 1 - \frac{\text{Levenshtein\_Distance}(s_1, s_2)}{\max(\text{length}(s_1), \text{length}(s_2))} \tag{2}$$

where:

- $s_1$ and $s_2$ are the strings being compared.
- Levenshtein_Distance$(s_1, s_2)$ calculates the minimum number of single-character edits (insertions, deletions, or substitutions) required to change one string into the other.
- length$(s)$ represents the length of string $s$.

Table 2 shows a list of some definitions that were extracted from the dataset. The rest of definitions can be found on the GitHub repository.[6]

### 3.2. Digital twin components identification

As shown in the examples in Table 2, each definition is comprised of a set of key components. For instance, the first one highlights components like "multi-physics," "multi-scale," "probabilistic simulation," "physical model," and "sensors." Meanwhile, the next example introduces other elements such as "software representation," "predict and optimize," and "real-time analytics".

To systematically identify the components of digital twins, we first conducted a Frequency Distributions Analysis from literature and experts, this process involves counting the occurrences of terms and phrases associated with digital twins in the extracted definitions dataset and within the survey dataset. The analysis provides numerical data on the most frequently referenced components and concepts, highlighting
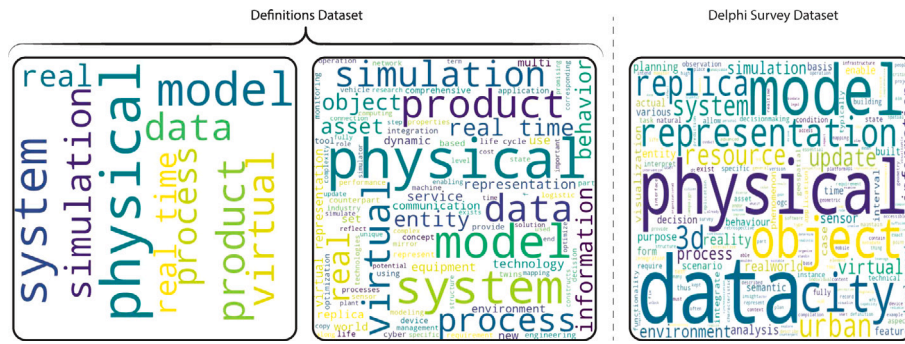
---

6 https://github.com/ualsg/dt-definitions

**Fig. 7.** Word cloud visualization of word frequency analysis of the extracted definitions dataset (on the left) and the expert survey (on the right).

**Table 3**
Generalized digital twin components.

| Category | Description | Related component |
|---|---|---|
| Data | Data representations of both the target entity and the digital entity. | 2D and 3D data<br>Real-time data<br>Data modeling |
| Analysis and services | Setup for modeling and analysis toward specific services, application use case and scenarios. | Simulation models<br>Data analytics and AI/ML models<br>Data Catalogue |
| Infrastructure | Software and hardware components, and other relevant tools supporting the operation of the digital twin system. | Cloud platform and architecture<br>High-Performance Computing (HPC)<br>Internet of Things (IoT) and sensor network |
| Interface | Requirements and setup for connection and interaction among digital twin components, as well as with users. | Application Programming Interface (API)<br>Visualization<br>Dashboards |
| System governance | System and design characteristics of the digital twin, relating to integration, governance, security and privacy, and ethics. | Data validation<br>Security protocols<br>Policy (and ethical protocols)<br>User management and administration |

prevalent elements and revealing emerging trends and gaps in current digital twin architectures. Fig. 7 illustrates the word cloud frequency analysis from the extracted definitions (on the left) and the expert survey (on the right).

We can observe that the top words identified in the frequency analysis in the definitions dataset closely correspond to the frequency analysis within the expert survey dataset, with a few differences, as seen in Fig. 7. For example, while 'physical', 'model', and 'data' are frequently mentioned in both the research articles and the survey, experts put more focus on '3D' and 'update' and researchers highlight the function of 'simulation' in DTs. These differences reflect the influence of varying perspectives across domains. However, an equivalence can be seen, considering the ideas of a DT implementation, for example 'update' from the expert survey could relate to the idea of 'real-time', or 'virtual' from the research articles may be synonymous to the 'representation' that experts refer to.

We can further note that the top words identified have agreement and are closely captured within the ISO 30173 DT system elements descriptions and DUET architecture specifications described in Section 2. This correspondence of terms allows us to identify generalized DT components for conducting further analysis on the dataset. We refer to top words identified from the frequency analysis word cloud and unify with related components as described in the DUET system architecture, ISO 30173 and Five-dimensional reference model to create a list of representative DT components, grouped into 5 generalized categories as summarized in Table 3, and described in more detail in the continuation.

1. **Data**

   DTs are envisioned to deal with multi-temporal and -scale heterogeneous data, which could be from multiple sources, either obtained from physical entities (e.g., as attribute data or dynamic condition data from sensors) or generated by virtual models (e.g., results from simulations within the DT) in both directions (Bi-directional) either from the physical twin to the virtual twin or vice versa using actuators or using human in the loop. We generalize three related components under this category:

   - *2D and 3D data:* The data (e.g., physical geometries, properties, behaviors, and rules) corresponds to either the target entity or digital entity.
   - *Real-time data:* Real-time data updates (i.e., of the target entity as facilitated by sensors and Internet of Things) or dynamic data outputs from real-time simulation process. Also, this includes data that are transferred the other direction, i.e, from the virtual twin to the physical twin via actuators or human-informed-decision in the loop.
   - *Data modeling:* Formalized and standardized digital representation of data, including definitions for data models, schema, ontologies, and digitization methods.

2. **Analysis and services**

   In specific application domains, the DT system may be designed with capabilities to provide specific services (e.g., integration,
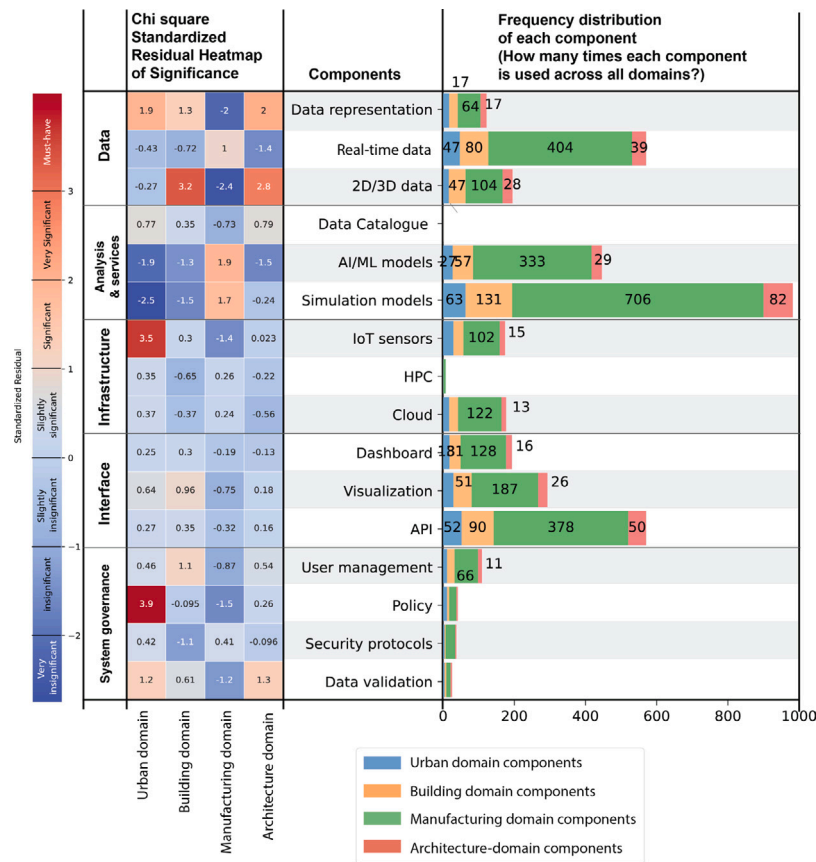
**Fig. 8.** Chi-square frequency distribution heatmap across different domains and component frequency distribution across different domains.

analysis, simulation, visualization, and optimization). We use the three general representative components for this category as follows:

- *Simulation models:* Implemented modeling of the specific processes (or scenarios) for representing the target entity or system.
- *Data analytics and AI/ML models:* Functionalities of the DT for running data analytics, AI, and ML algorithms and models, i.e., implemented for specific use cases or services.
- *Data Catalogue:* Relate to DT data management, including setup for data discovery and data access, e.g., knowledge graphs and context graphs.

3. **Infrastructure**

Infrastructure is a major component supporting the DT to ensure its normal operation, which could include a slew of different hardware components as required for specific use cases. We include the three below for our analysis:

- *Cloud platform and architecture:* The technological components specifically set up for the DT in utilizing cloud environments for services.
- *High-Performance Computing (HPC):* DT component for supercomputing, i.e., to handle huge volumes of data and complex models at high speeds.
- *Internet of Things (IoT) and sensor network:* The network of physical devices interconnected and communicating with each other for the collection and sharing of data.

4. **Interface**

Within the DT system, digital entities are connected dynamically with their physical counterparts. Data connections enable data exchange between entities and link the DT's models and services. Further, the DT is expected to have a platform for the user to get information and interact with the data and models.

- *Application Programming Interfaces (APIs):* Software implementation to facilitate the access and control of data flow between components (i.e., as data gateways), as well as between DT and user (i.e., through dashboards or other apps).
- *Visualization:* The functionality of the DT for providing data presentation and other visual outputs.
- *Dashboards:* The platform (and tools) for data and visualization display, and for DT user interface (i.e., through web applications) that facilitate the users' access to the DT data, models and outputs.

5. **System governance**

Aspects of integration, governance, security and privacy, and data ethics are requirements for a DT system to coexist with other DT systems) [104], for example, matters related to data collection and access. Such system aspects could include (but are not limited to) the following:

- *Data validation:* Refer to control measures for data checks to ensure the meeting of required data quality and accuracy.
- *Security protocols:* Protocols specific to the prevention of risks (e.g., hacks, data leaks, etc.) and securing of the DT's operation.

- *Policy:* Includes policies in place to ensure ethical use and management of the DT data and products.
- *User management and administration:* Protocols and provisions for implementing user authorization and access, consistent with security and policy.

### 3.3. Component identification and analysis

We conducted component frequency and Chi-square analysis to examine the association between DT components (Table 3) and various domains, including urban, building, manufacturing, and architecture. The use of Chi-square test is particularly useful in our study due its ability to handle categorical data. The significance was tested against the null hypothesis $H_0$ (No significant contribution) and the alternative hypothesis $H_1$ (There is a significant contribution) for each component. The chi-square statistic of 109.76 and a *p*-value of 2.48e−07 indicated a significant relationship between the components and their respective domains. We calculated the Standardized residuals to identify specific deviations, with significant residuals (absolute values greater than 2) highlighting components that were either over-represented or under-represented. Notably, "IoT sensors" and "Policy" were significantly more frequent in the urban domain, while "2D/3D data" showed higher frequencies in the building and architecture domains but lower in manufacturing. This analysis, including adjustments for zero and low values, provided a detailed understanding of the distribution and significance of digital twin components across different fields, offering critical insights for the development and application of DT in various sectors (Figs. 8). We also performed a correlation and dendrogram clustering analysis between all component residuals to explore which components are joined in terms of significance and insignificance across all domains (Fig. 10). The analysis clearly divides these components into two big clusters of correlations.

To extract the definitions across domains, we compared the definitions from the existing dataset (800 definitions) and the expert survey to the components analysis for each domain. LLM was used to extract the components from these definitions and compare them with the components identified through frequency analysis.

## 4. Results

### 4.1. Component analysis across domains

Based on the Chi-square test and the correlation clustering of components, it is noticed that Cloud computing, HPC, Security protocols, Real time data, simulation models, and AI/ML models are highly correlated with each other and can be segregated into one group of components. On the other hand, Data representation, Data validation, Visualization, and Policy are highly correlated with each other and can be segregated into another group of components. We refer to the first group of components as HPRT (High-Performance Real-Time) Digital Twins, characterized by real-time, big-data-driven, dynamic, and adaptable components. The second group is referred to as LTDS (Long-Term Decision Support) Digital Twins, characterized by near-to-far real time data management and visualization components. The first group of components is used in computationally intensive applications, such as manufacturing, self-driving cars, and aerospace and rarely used in the BE and AEC domains. In contrast, the second group is suited for applications requiring long-term decision support, such as urban planning and building management.

The temporal analysis further highlights the evolving nature of these components (Fig. 9). While some components like AI/ML and real-time data capabilities have seen rapid growth in adoption due to advancements in technology and increasing data availability, others like policy and data representation have remained consistently significant due to their foundational role in ensuring the accuracy and applicability of digital twins in long-term planning and management contexts.

The temporal evolution also shows some interesting insights about the adoption of some components; for example, Data Catalogues, Dashboards, and IoT sensors are lagging behind others, such as Data representations, API, Visualizations, and Systems as a result of the growing awareness of data issues such as data heterogeneity and quality [128].

In the following sub-sections, we provide comprehensive analytical insight and key results focusing on each domain.[7]

### Building domain

Components like 2D/3D data (residual of 3.2), data representation (1.3), and visualization (0.96) are significantly more frequent than expected, underscoring their critical importance in accurately modeling and managing building structures in the form of Building Information Models (BIM) [129–132] (Fig. 11 - a).

On the other hand, a negative residual indicates that simulation models are less frequent in the building domain than expected. This might suggest that while simulation models are important, they may not be as widely implemented in building digital twins as in other domains. Notably, real-time data, AI/ML, Security protocols, and HPC are not mature enough in the building digital twin domains in research and practice due to technical, organizational, and data-related aspects [30,133]. For instance, the integration of BIM and real-time data streams from IoT sensors is complex [134], interoperability between different data schemas, sources, software and platforms is difficult, in addition to scalability issues [130,135,136].

### Architectural domain

Architecture terms show significance values closer to the building terms. Digital Twins are used in Architecture for various purposes, such as design and planning, construction management, and operation and maintenance. The architecture domain is characterized by the prominent use of 2D/3D data (Residual of 2.8), data representation (residual of 2), and validation (residual of 1.3) (Fig. 11 - b).

The negative residual of simulation models indicates that simulation models are less frequent in the architecture domain of digital twins than expected. This finding might suggest that while simulation models are important, they may not be as widely implemented in architecture digital twins as in other domains.

AI/ML models (residual of −1.5) and Real-time data (residual of −1.4) are less significant in the architecture domain. This might be due to the lack of data and the complexity of the data in the architecture domain.

Despite these challenges, the use of digital twins in architecture is growing. Architects are increasingly leveraging digital twins for improved project visualization, enhanced collaboration among stakeholders, better lifecycle management of buildings, facility management [137], and construction monitoring [138,139]. For instance, digital twins can simulate the impact of design changes on building performance, enabling architects to make more informed decisions. Furthermore, the integration of augmented reality (AR) and virtual reality (VR) with digital twins is opening new avenues for immersive design experiences and client presentations. As technology continues to advance, it is expected that the adoption of AI/ML models and real-time data in architectural digital twins will increase, providing more sophisticated tools for predictive analytics and real-time monitoring of building systems.

---

[7] Although Building, Architecture, and Urban terms can be identified as different scales of the BE or the AEC rather than domains perse, they are segregated into different domains in this study because it is noted that there are some differences in the requirements and understanding of DT of each scale.
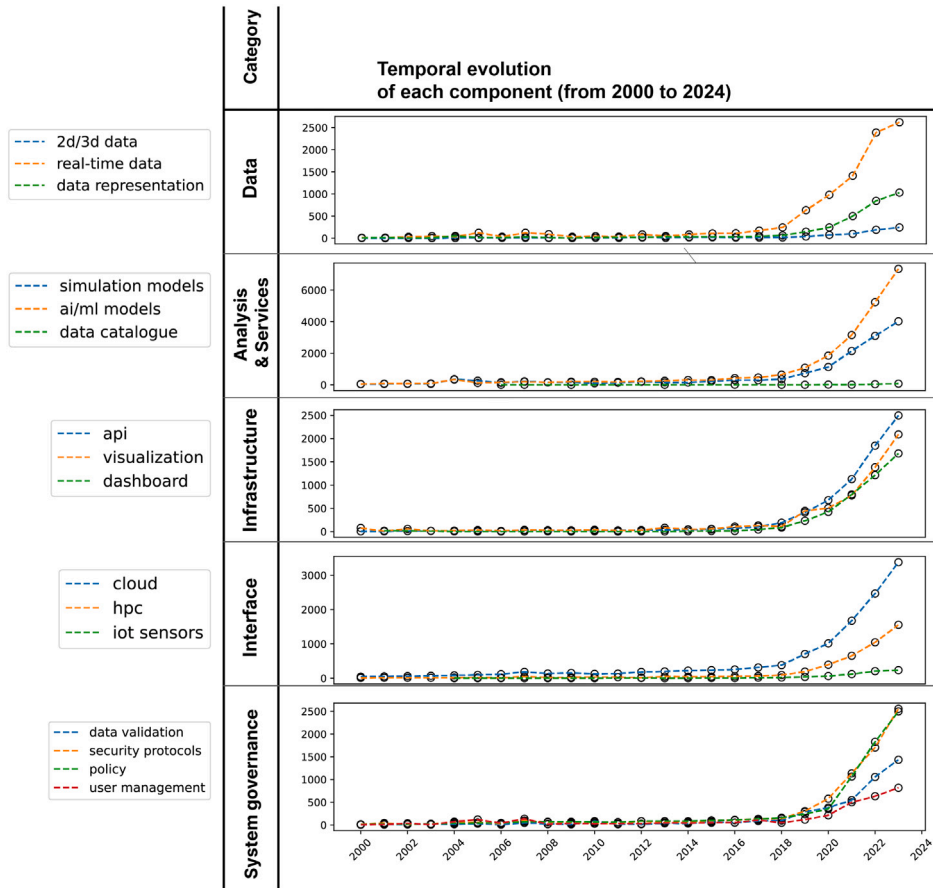
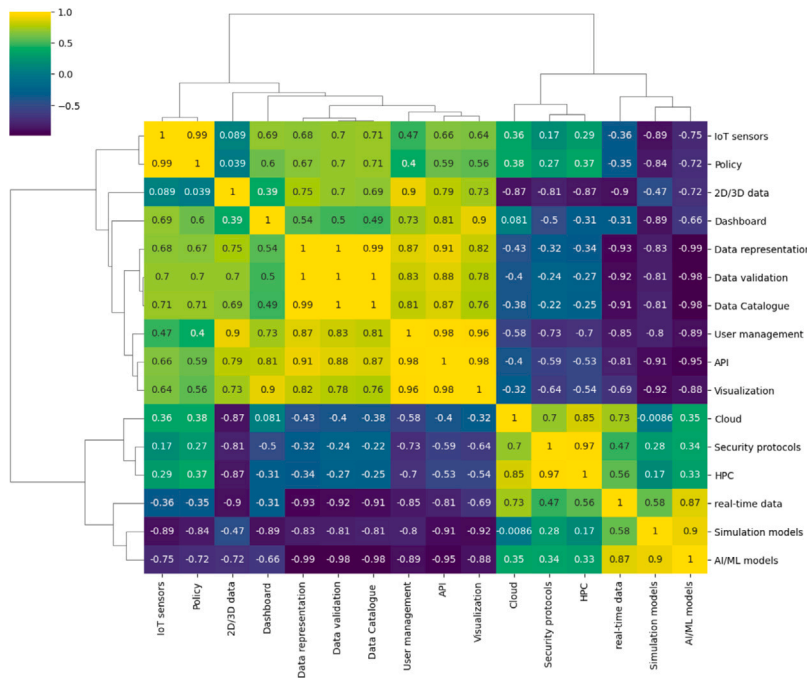**Fig. 9.** Temporal evolution of components from 2000 to 2024.



**Fig. 10.** Clustered correlation between components residuals across all the domains. The higher correlation between two components means that these two components have either significant or non-significant residual.
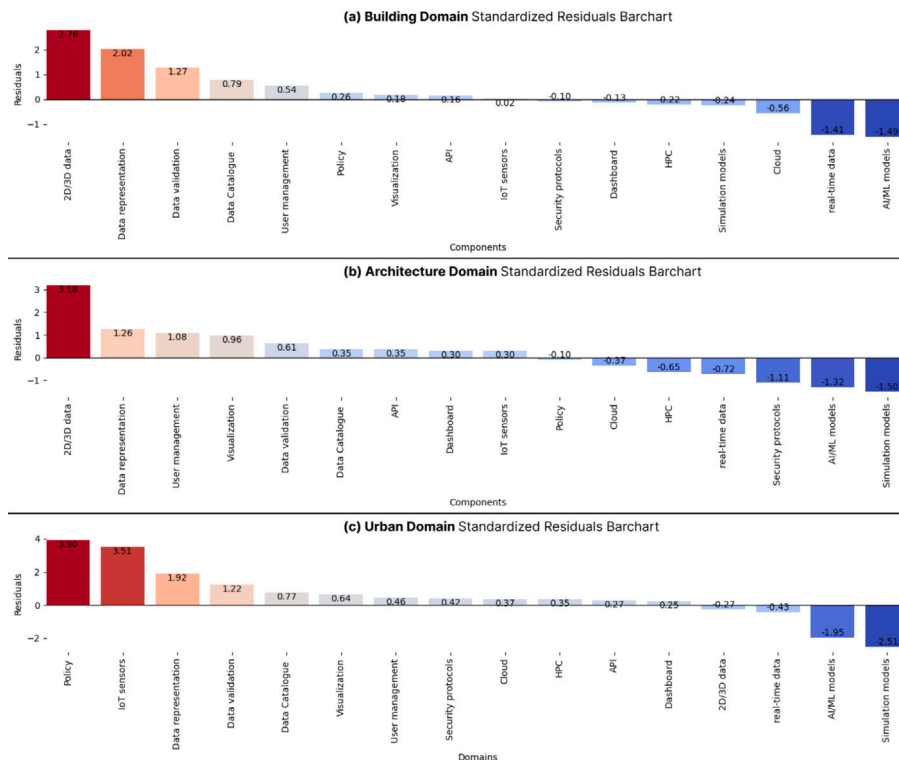
**Fig. 11.** Components for the BE domains (a) Building, (b) Architecture, and (c) Urban sorted by significance. The manufacturing domain is excluded since its results is out of the scope of the research.

*Urban domain*

Urban terms in digital twins (UDTs) reveal various insights and challenges. Our analysis shows that there are fewer significant components in UDT given the complexity of the domain and the requirements of massive computational and storage resources, among other challenges that hindered the adoption of some components, such as real-time data and AI/ML in the urban domain. In contrast, IoT sensors, data representation, data validation, and policy-driven decision-making show greater significance (Fig. 11 - c).

The dynamic and transient nature of cities requires computational resources and expertise across various domains. Acquiring big amounts of real-time data, primarily through sensors, is essential for enabling Urban Digital Twin (UDT). Despite the increasing number volume of data captured through IoT sensors (Residual of 3.5), many researchers argue that the sensor network in cities is still too sparse and heterogeneous [140]. This introduces significant challenges related to data acquisition, synchronization, and applications, which limited its use in UDTs.

The sparse and heterogeneous nature of IoT sensor networks affects AI/ML applications in urban digital twins. **AI/ML models (Residual of −1.9)** show less significance in the urban domain than expected, as these applications require massive computational power to handle big data from various sources. Several researchers argue that there is not enough data for proper AI/ML applications in urban digital twins. However, the trend (Fig. 9) indicates that the adoption of ML and AI in urban digital twins has been rapidly growing and gaining significant traction since 2020 [141,142]. This growth is driven by advancements in computational technologies and the increasing availability of urban data from various resources, facilitating more sophisticated and scalable AI/ML applications in urban analytics and planning. A recent advancement in this field is GeoAI [143–146], which integrates AI/ML and data science with geoscience to enhance the capabilities of urban studies, including digital twins [43].

**Simulation models (Residual of −2.5)** indicate less frequent adoption of Simulation models in urban and city digital twin contexts. This

is attributed to the dynamic and transient nature of cities, which requires adaptable models and hence substantial computational resources and expertise across various domains [147]. Many of these existing simulation models on the city scale are characterized by oversimplification [148].

Decision-making-enabled digital twins play a key role in the success of digital twin models, appearing significant in terms of **Policy within UDTs (Residual of 3.9)**. Urban policy typically focuses on the causal dynamics of data insights to address issues such as housing, transportation, public health, economic development, and social equity. Policies establish guidelines for the collection, management, and protection of big data and guide the strategic application of urban analytics to address long-term structural challenges facing cities [140].

From the component analysis, we conclude that Building, Architecture, and Urban DTs can be categorized as long-term decision support rather than high-performance real-time systems. We acknowledge that there are several attempts to adopt real-time digital twin applications on city scales, however, there is no real and widely prominent HPRT scalable DT on city/urban scales to the best of the authors' knowledge. Thus, we categorize BDTs, ADTs, and UDTs as follows:

- **Building Digital Twins**: These digital twins emphasize accurate modeling, data representation, and visualization to support building management and operations. Key components include 2D/3D data, data representation, and visualization. Applications involve Building Information Models (BIM), building management systems, and lifecycle management of building structures. They help in accurate modeling, managing structures, and ensuring long-term operational efficiency.
- **Architectural Digital Twins**: Similar to building digital twins, these are focused on long-term planning and decision support, leveraging detailed data representation and validation. Key components include 2D/3D data, data representation, and validation.
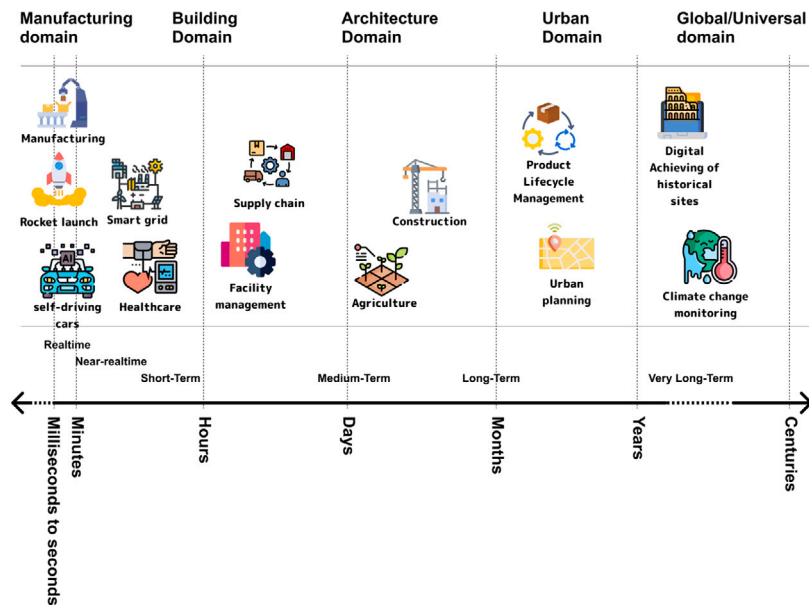
**Fig. 12.** Applications of Digital Twins from real-time to very-long term temporal scales, suggesting that building, architecture, and urban DTs do not necessarily require real-time response as much as other domains such as manufacturing.
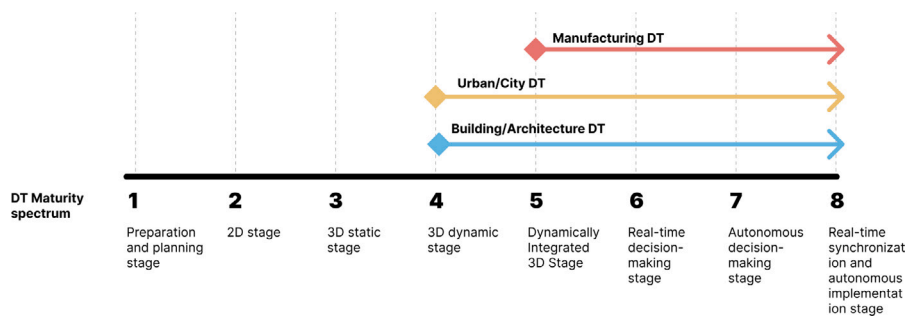


**Fig. 13.** Approximate digital Twin maturity for each domain.

Applications include design and planning, construction management, operation, and maintenance. They are used to improve project visualization, enhance stakeholder collaboration, and manage building lifecycles.

- **Urban Digital Twins (UDTs)**: **LTDS** (Long Term Decision Support) Digital Twins. These digital twins focus on data representation, data validation, and policy-driven decision-making. Key components include IoT sensors, data representation, data validation, and policy. Applications involve urban planning, public health, transportation management, economic development, and social equity. These applications require long-term planning and policy formulation supported by data insights.

### 4.2. Derived definitions and implications

In this section, we show the results of the derived definitions based on the frequency analysis. The results reveal some divergences from the generalized definitions of digital twins. For example, when focusing on the urban/city domains, real-time data, and AI/ML models are not a significant portion of the definitions (Fig. 12). Also, given our analysis results, we identified minimum level of maturity requirements for each domain based on the CITYSTEPS framework by [46].

The following subsections present the findings from the LLM analysis of the dataset, highlighting the top qualifying definitions that appeared in the literature for each domain and our derived definitions.

#### 4.2.1. Building and architecture digital twin (BDT and ADT) definitions

From the existing dataset, we extracted definitions that focus on comprehensive data representation, lifecycle management, and the ability to simulate and optimize operations without restrictions on data update frequencies. Such definitions are well-aligned with the significance test for building digital twins. These elements ensure that the digital twin can provide high-value services and support long-term asset management and operational efficiency:

"**A digital twin is a virtual representation of a physical asset or system that securely holds all relevant static and dynamic information from concept to decommissioning, enabling high-value collaborative services**" by [149].

This definition comprehensively emphasizes that a digital twin is a virtual representation that holds all relevant static and dynamic information, including 2D/3D data, data representation, data visualization, and data cataloging. Additionally, it mentions "enabling high-value collaborative services" which incorporates user management. The definition also does not restrict the frequency of data collection, processing, and action taking which aligns with the building DTs significance test. On the other hand, the definition focuses on the importance of the lifecycle of the asset/system, which spans over short to long-term frequencies and updates.

Another definition is "**Digital twin is a living model of the physical asset or system, which continually adapts to operational changes based on the collected online data and information and can forecast the future of the corresponding physical counterpart**"

by [150]. Although the definition mentions continual adaptation of changes, it does not restrict this adaptation to real-time. Furthermore, the reference to "... can predict the future" implies that ML/AI and simulation are considered "good-to-have" components rather than must-have components in a DT system.

A third definition that is also aligned with the significance test is the following: **"Digital twin is a virtual replica of a physical asset, such as a dam, that can be used to simulate and test various scenarios, monitor performance, and optimize operations"** as mentioned by [151]. Although this definition gives an example of dams, it can also be applicable to building and construction. It does not have restrictions on the time-interval data updates between the physical and the virtual twins. Additionally, it highlights the capability to simulate and test various scenarios, which is crucial for optimizing performance and operations in building DTs. The flexibility in application and emphasis on optimization and monitoring are key aspects that resonate with the significance test.

A more generic building-oriented definition is: **"The digital twin is a digital representation of an asset or system and mimics its real-world behavior with the aim of managing, planning, predicting, and demonstrating current and future construction processes and models via synchronized representations"** by [152]. The author also emphasized that ... "The digital twin can capture overall changing site conditions in near real-time and creates a virtual simulation environment ..." which indicates that the level of detail and time–frequency is subject to the nature of the asset, system, and the application of the digital twin. In this context, the authors refer to **"synchronized representations"**.

Another definition that leans more toward a generic digital twin but also aligns with the Building DT components: **"Digital Twin is defined as a software representation of a physical asset, system, or process designed to detect, prevent, predict, and optimize through analytics to deliver business value"** by [153].

From the extracted definitions of the BDT, we have deduced the following definition:

**Definition 1. A Building/Architecture Digital Twin (BDT/ADT)** is a spatial–temporal virtual representation of a physical asset or system that manages all relevant static and dynamic information. It enables collaborative services, adapts to operational changes at suitable intervals, and can forecast future scenarios for optimized performance and decision-making.

Gathering insights from the experts' responses, we highlighted "accurate 3D geometry of buildings" and "rich semantic information" as critical components in the BE DTs. Such findings align with the extract definitions from the developed corpus of research articles. According to the CITYSTEPS framework, for a BDT/ADT to be considered a true digital twin, it must reach at least maturity level 4 i.e., the 3D dynamic stage (Fig. 13). This dynamic aspect can range from real-time updates to long-term changes (Fig. 12). Therefore, a BIM model, whether hosted online or on local servers, that lacks regular updates reflecting changes in the real-world entity does not meet the criteria for a digital twin.

*4.2.2. Urban and city digital twin (UDTs and CDTs) definitions*

In the urban domain, there is no specific definition that accurately addresses all the components and their significance. However, some of these definitions are sufficiently generic to be applicable to urban/city/geospatial contexts. For example, one definition states **"Digital Twin is a dynamic representation of a real-life object that mirrors its states and behavior across its lifecycle and that can be used to monitor, analyze and simulate current and future states of and interventions on these objects, using data integration, artificial intelligence, and machine learning."** [74,154]. This definition refers to mirroring the state and behavior of an object across its lifecycle. Although this could encompass that the mirror is in real-time. The

authors mentioned that it is in near real-time. However, in the city and urban context, mirroring behavior does not necessarily need to be momentary; it can occur over various time spans (real-time, near real-time, short, mid, or long term intervals). Extensive data updates may lead to redundant and unnecessary information, computational and storage burdens, and, in many cases, may not inform decision-making effectively.

Another definition, the one by [155], states that **"digital twin is a virtual model that replicates a real-world urban area/city, allowing the creator to determine the performance of its different systems"**. Although this definition is explicitly dedicated to the urban/city domain, it does not include information about the contribution of the DT in defining policies or informing decision making, which is a significant part of UDTs and CDTs.

Evaluating the findings from our expert survey, we observe an alignment between research articles and experts' responses. For example, one expert defined an urban digital twin as "a common and shareable (data) description of the current (physical) reality of the city". However, several experts indicate critical components of digital twins in the survey which indeed contradict the analysis of systematic review. For example, one definition refers a digital twin to "a general term for digital (virtual) replication of actual entities (object/processes), updated from real-time data, for supporting decision-making". Further, some practitioners do not consider a 3D city model "without automatic update procedures" nor "not reflecting real-time situation" as a digital twin. Notably, the component analysis suggests that real-time data updates may not be necessary in the urban domain, given that urban development takes a relatively long time. Nevertheless, one expert addressed that 'no comprehensive definition of the term urban digital twin has been provided yet'.

Based on the information extracted from the definitions, component analysis, and the expert survey, we derive the following definition for Urban and City Digital Twins (UDT and CDT).

**Definition 2. A City/Urban Digital Twin** is a spatial–temporal virtual representation of a real-world urban area or city, mirroring its states during its lifecycle through IoT sensors. It is used to monitor and analyze urban systems across different time spans to aid in decision-making and can be extended to simulate and predict various states and scenarios.

In this definition, we emphasized that the scale of an UDT is a city or an urban area, and it must be a replica of an existing physical setting. IoT sensors is a key component in an urban digital twin due to the complexity and dynamics of cities. Data representation, monitoring, and analysis are a significant part of a CDT. Also, these data are used to make decisions in different time spans. On the other hand, a CDT can be extended to include real-time data integration, AI/ML, and simulation models. However, these components are not essential. According to the CITYSTEPS framework (Fig. 13) for digital twin maturity, a CDT/UDT must reach at least level 4 (the 3D dynamic stage) to qualify as a digital twin. As with BDT/ADT, there is no strict requirement for the frequency of dynamic updates, which can range from real-time to long-term intervals or anything in between.

**5. Discussion and future directions**

*Comparison with traditional approaches*

Traditional approaches to analyzing definitions and concepts, such as manual literature reviews or expert-driven content analysis, often rely on smaller datasets due to the time-intensive nature of these methods. While these methods can provide nuanced insights, they are inherently limited in scalability and subject to potential bias introduced by individual reviewers.

Our approach leverages advanced NLP techniques, enabling us to process and analyze a large corpus of 15,000 articles efficiently. This methodology provides distinct advantages over traditional methods.

A key strength of this approach lies in its ability to derive definitions from a collective understanding of the concept. By analyzing a large and diverse set of papers discussing DTs across different disciplines, our methodology captures the broader agreement within the scientific community while identifying nuances and variations. This collective perspective enables the development of a definition that reflects a consensus-driven understanding, ensuring relevance and applicability across domains.

*Current limitations in the BE DTs*

Our study has identified several limitations in the current interpretation and implementation of DTs within the BE. One significant limitation is the lack of real-time data integration and advanced AI/ML models despite being recognized by the majority of experts as essential. Our analysis and review showed that the adoption of such components is minimal. This gap is partly due to the complexity of integrating heterogeneous data sources and the high computational demands required for real-time processing on buildings and city scales.

*The future of ML/AI in BE DTs*

The integration of ML and AI in BE DTs is predicted to revolutionize the field by enabling predictive analytics, automated decision-making (Stage 7 on the CITYSTEPS model of maturity), and enhanced simulation capabilities. Future research should focus on developing scalable AI/ML models capable of handling the vast amounts of data generated in BE applications. This includes improving data quality and developing robust algorithms that can operate efficiently in real time.

*Multi-scale DTs (from human scale to the global scale)*

*One size does not fit all* when it comes to a comprehensive definition of DTs. Although there are a few common denominator components of any DT, different scales (human, building, district, urban, city, global, and universal) DTs have different components. The organization and priority of these components differ significantly based on the use case. Future research should explore the development of multi-scale DT frameworks that seamlessly integrate data and insights across different scales, enabling a holistic approach to BE management.

*Fully autonomous BE DTs*

By tracking the temporal trends and the evolution of DT components over the years, a clear trajectory toward fully autonomous DTs can be observed (Stage 8 of Maturity on the CITYSTEPS scale). The advancements in AI/ML, real-time data integration, and decision-making algorithms are key drivers in this direction. Autonomous DTs would possess the capability to continuously monitor and adapt to changes in the environment without human intervention, providing real-time responses to emerging issues and optimizing operations based on predictive insights.

Achieving full autonomy in BE DTs necessitates overcoming significant technical challenges, including the integration of advanced AI/ML models capable of handling vast and diverse datasets, ensuring robust data security and privacy, and developing seamless interoperability between different systems.

*Generative models for DTs*

Generative models hold great potential for enhancing the capabilities of DTs. These models can create realistic simulations of various scenarios, providing valuable insights for decision-making and planning. By leveraging generative models, DTs can simulate the impact of different design choices, environmental changes, and operational strategies. Future research should focus on integrating generative models with DT frameworks, enabling more accurate and comprehensive simulations.

## 6. Conclusion

Digital twin is now a common concept and technology in disciplines such as building science and urban planning. But its meaning is contested and often subject to specific interpretations, and there is a degree of gerrymandering as well. In this research, we aspired to make sense of the terminology. We developed a novel method to do so, conducted an analysis of a massive set of descriptions of Digital Twin in scientific literature within the built environment using NLP, LLMs, and statistical approaches, capturing around one thousand definitions that appeared in the literature from 2000 to 2024. We extracted 15,000 full-text articles about digital twins and built on top of a pre-deployed Delphi expert survey [26] involving 52 experts in the related domains. Our goal was to identify the core components and characteristics of digital twins, focusing on the urban, architectural, and building domains. We depended on extracting DT components from a large corpus of text and then assessing these components against the standardized definitions and the existing definitions from other domains (e.g., Manufacturing). We conducted a frequency analysis and a Chi-Square statistical significance test to evaluate the prominence of various DT elements in specific domains.

The findings revealed significant variations in digital twin descriptions and paradigms across different domains, highlighting the complexity and evolving nature of this concept. We identified two primary categories of digital twins, Long Term Decision Support (LTDS) DTs and High-Performance Real Time (HPRT) DTs. We found that all the built-environment-related DTs fit on the first one.

Although this research addresses the current definition of Digital Twins (DTs), future trends indicate that some components are becoming increasingly prominent such as AI/ML, HPC, and Real-time data adoption. Thus, future research is needed to redefine DTs according to new trends, techniques, and tools. This reflects the fact that Building, Urban, and City DTs have not yet reached their state of equilibrium regarding a fully indicative definition due to their complexity and ever-changing landscapes. This research offers a foundational framework for defining Digital Twins in the built environment and serves as a first step toward building consensus. The methodology and definitions presented here provide a sustainable approach that encourages researchers and practitioners to adopt these definitions in their implementations or refine them based on specific requirements. We encourage researchers to deploy the same methodology and revisit the definition periodically to track its evolution over time, reflecting advancements in technology and evolving understandings of the DT entity.

**CRediT authorship contribution statement**

**Mahmoud Abdelrahman:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Edgardo Macatulad:** Writing – original draft, Validation, Methodology, Investigation, Funding acquisition. **Binyu Lei:** Writing – review & editing, Validation, Investigation, Formal analysis, Data curation. **Matias Quintana:** Writing – review & editing, Visualization, Software, Formal analysis. **Clayton Miller:** Writing – review & editing, Conceptualization. **Filip Biljecki:** Writing – review & editing, Visualization, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

## Declaration of Generative AI and AI-assisted technologies in the writing process

## Funding

## Declaration of competing interest

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

## Acknowledgments

## Data availability

The authors do not have permission to share data.

## References

[1] D. Gelernter, Mirror Worlds: Or the Day Software Puts the Universe in a Shoebox. How It Will Happen and What It Will Mean, Oxford University Press, 1993.

[2] M.W. Grieves, Product lifecycle management: the new paradigm for enterprises, Int. J. Prod. Dev. 2 (2005) 71–84.

[3] J. Leng, D. Wang, W. Shen, X. Li, Q. Liu, X. Chen, Digital twins-based smart manufacturing system design in Industry 4.0: A review, J. Manuf. Syst. 60 (2021) 119–137, http://dx.doi.org/10.1016/j.jmsy.2021.05.011.

[4] Y. Jiang, S. Yin, K. Li, H. Luo, O. Kaynak, Industrial applications of digital twins, Phil. Trans. R. Soc. A 379 (2021) 20200360, http://dx.doi.org/10.1098/rsta.2020.0360.

[5] H. Lehner, L. Dorffner, Digital geotwin vienna: Towards a digital twin city as geodata hub, 2020, http://dx.doi.org/10.1007/s41064-020-00101-4.

[6] F. Dembski, U. Wössner, M. Letzgus, M. Ruddat, C. Yamu, Urban digital twins for smart cities and citizens: The case study of Herrenberg, Germany, Sustainability 12 (2020) 2307, http://dx.doi.org/10.3390/su12062307.

[7] M.G. Juarez, V.J. Botti, A.S. Giret, Digital twins: Review and challenges, J. Comput. Inf. Sci. Eng. 21 (2021) http://dx.doi.org/10.1115/1.4050244.

[8] M. Sommer, J. Stjepandić, S. Stobrawa, M. von Soden, Automated generation of digital twin for a built environment using scan and object detection as input for production planning, J. Ind. Inf. Integr. 33 (2023) 100462, http://dx.doi.org/10.1016/j.jii.2023.100462, URL: https://www.sciencedirect.com/science/article/pii/S2452414X23000353.

[9] M. Bartos, B. Kerkez, Pipedream: An interactive digital twin model for natural and urban drainage systems, Environ. Model. Softw. 144 (2021) 105120.

[10] I. Chomiak-Orsa, K. Hauke, K. Perechuda, M. Pondel, The use of digital twin in the sustainable development of the city on the example of managing parking resources, Procedia Comput. Sci. 225 (2023) 2183–2193.

[11] G. White, A. Zink, L. Codecá, S. Clarke, A digital twin smart city for citizen feedback, Cities 110 (2021) 103064.

[12] P. Alva, F. Biljecki, R. Stouffs, Use cases for district-scale urban digital twins. The international archives of the photogrammetry, Remote. Sens. Spat. Inf. Sci. XLVIII-4/W4-2022 (2022) 5–12, http://dx.doi.org/10.5194/isprs-archives-XLVIII-4-W4-2022-5-2022.

[13] P. Liu, T. Zhao, J. Luo, B. Lei, M. Frei, C. Miller, F. Biljecki, Towards human-centric digital twins: Leveraging computer vision and graph models to predict outdoor comfort, Sustain. Cities Soc. 93 (2023) 104480, http://dx.doi.org/10.1016/j.scs.2023.104480.

[14] I. Jeddoub, G.A. Nys, R. Hajji, R. Billen, Digital twins for cities: Analyzing the gap between concepts and current implementations with a specific focus on data integration, Int. J. Appl. Earth Obs. Geoinf. 122 (2023) 103440, http://dx.doi.org/10.1016/j.jag.2023.103440.

[15] J. Metcalfe, C. Ellul, J. Morley, J. Stoter, Characterizing the role of geospatial science in digital twins, ISPRS Int. J. Geo- Inf. 13 (2024) 320, http://dx.doi.org/10.3390/ijgi13090320.

[16] S. Shirowzhan, W. Tan, S.M.E. Sepasgozar, Digital twin and cybergis for improving connectivity and measuring the impact of infrastructure construction planning in smart cities, ISPRS Int. J. Geo- Inf. 9 (2020) 240, http://dx.doi.org/10.3390/ijgi9040240.

[17] J. Yan, S. Zlatanova, M. Aleksandrov, A.A. Diakite, C. Pettit, Integration of 3D objects and terrain for 3D modelling supporting the digital twin. ISPRS annals of the photogrammetry, Remote. Sens. Spat. Inf. Sci. IV- 4/W8 (2019) 147–154, http://dx.doi.org/10.5194/isprs-annals-iv-4-w8-147-2019.

[18] A. Belfadel, S. Hörl, R.J. Tapia, D. Politaki, I. Kureshi, L. Tavasszy, J. Puchinger, A conceptual digital twin framework for city logistics, Comput. Environ. Urban Syst. 103 (2023) http://dx.doi.org/10.1016/j.compenvurbsys.2023.101989, URL: https://www.sciencedirect.com/science/article/pii/S0198971523000522.

[19] M.G. Kim, M. Bartos, A digital twin model for contaminant fate and transport in urban and natural drainage networks with online state estimation, Environ. Model. Softw. 171 (2024) 105868.

[20] S. Wang, L.H. Vu, The integration of digital twin and serious game framework for new normal virtual urban exploration and social interaction, J. Urban Manag. 12 (2023) http://dx.doi.org/10.1016/j.jum.2023.03.001, URL: https://www.sciencedirect.com/science/article/pii/S2226585623000225.

[21] B. Lei, Y. Su, F. Biljecki, Humans as sensors in urban digital twins, in: International 3D GeoInfo Conference, Springer, 2023c, pp. 693–706.

[22] B. Lei, X. Liang, F. Biljecki, Integrating human perception in 3D city models and urban digital twins, ISPRS Ann. Photogramm. Remote. Sens. Spat. Inf. Sci. 10 (2024) 211–218.

[23] J. Shi, Z. Pan, L. Jiang, X. Zhai, An ontology-based methodology to establish city information model of digital twin city by merging BIM, GIS and IOT, Adv. Eng. Inform. 57 (2023) http://dx.doi.org/10.1016/j.aei.2023.102114, URL: https://www.sciencedirect.com/science/article/pii/S1474034623002422.

[24] H. Xia, Z. Liu, M. Efremochkina, X. Liu, C. Lin, Study on city digital twin technologies for sustainable smart city design: A review and bibliometric analysis of geographic information system and building information modeling integration, Sustain. Cities Soc. 84 (2022) 104009, http://dx.doi.org/10.1016/j.scs.2022.104009, URL: https://www.sciencedirect.com/science/article/pii/S2210670722003298.

[25] C. Boje, A. Álvaro José Hahn Menacho, E. Benetto, S. Kubicki, T. Schaubroeck, T.N. Gutiérrez, A framework using bim and digital twins in facilitating LCSA for buildings, J. Build. Eng. 76 (2023) http://dx.doi.org/10.1016/j.jobe.2023.107232, URL: https://www.sciencedirect.com/science/article/pii/S2352710223014122.

[26] B. Lei, P. Janssen, J. Stoter, F. Biljecki, Challenges of urban digital twins: A systematic review and a delphi expert survey, Autom. Constr. 147 (2023a) 104716.

[27] D.G.J. Opoku, S. Perera, R. Osei-Kyei, M. Rashidi, Digital twin application in the construction industry: A literature review, J. Build. Eng. 40 (2021) 102726, http://dx.doi.org/10.1016/j.jobe.2021.102726, URL: https://www.sciencedirect.com/science/article/pii/S2352710221005842.

[28] A.P. Pereira, M. Buzzo, I. Zimermann, F.H. Neto, H. Malgarezi, A descriptive 3D city information model built from infrastructure BIM: Capacity building as a strategy for implementation, Int. J. E- Plan. Res. 10 (2021) http://dx.doi.org/10.4018/IJEPR.20211001.oa9.

[29] A. Bolton, L. Butler, I. Dabson, M. Enzer, M. Evans, T. Fenemore, F. Harradence, E. Keaney, A. Kemp, A. Luck, et al., Gemini principles, 2018, http://dx.doi.org/10.17863/CAM.32260, URL: https://www.cdbb.cam.ac.uk/system/files/documents/TheGeminiPrinciples.pdf.

[30] I. Osadcha, A. Jurelionis, P. Fokaides, Geometric parameter updating in digital twin of built assets: A systematic literature review, J. Build. Eng. 73 (2023) 106704, http://dx.doi.org/10.1016/j.jobe.2023.106704, URL: https://www.sciencedirect.com/science/article/pii/S2352710223008835.

[31] H.Y. Chong, H.H. Hosamo, M.H. Hosamo, Digital twin technology for bridge maintenance using 3D laser scanning: A review, Adv. Civ. Eng. 2022 (2022) 2194949, http://dx.doi.org/10.1155/2022/2194949.

[32] W. Kritzinger, M. Karner, G. Traar, J. Henjes, W. Sihn, Digital twin in manufacturing: A categorical literature review and classification, IFAC- Pap. 51 (2018) 1016–1022, http://dx.doi.org/10.1016/j.ifacol.2018.08.474, URL: https://www.sciencedirect.com/science/article/pii/S2405896318316021, 16th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2018.

[33] L. Li, S. Aslam, A. Wileman, S. Perinpanayagam, Digital twin in aerospace industry: A gentle introduction, IEEE Access 10 (2022) 9543–9562, http://dx.doi.org/10.1109/ACCESS.2021.3136458.

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

[35] C. Cimino, E. Negri, L. Fumagalli, Review of digital twin applications in manufacturing, Comput. Ind. 113 (2019) 103130, http://dx.doi.org/10.1016/j.compind.2019.103130.

[36] D. Yang, H. Karimi, O. Kaynak, S. Yin, Developments of digital twin technologies in industrial, smart city and healthcare sectors: A survey, Complex Eng. Syst. 1 (2021) 3.

[37] G. Steindl, M. Stagl, L. Kasper, W. Kastner, R. Hofmann, Generic digital twin architecture for industrial energy systems, Appl. Sci. 10 (2020) http://dx.doi.org/10.3390/app10248903.

[38] D. Jones, C. Snider, A. Nassehi, J. Yon, B. Hicks, Characterising the digital twin: A systematic literature review, CIRP J. Manuf. Sci. Technol. 29 (2020) 36–52, http://dx.doi.org/10.1016/j.cirpj.2020.02.002, URL: https://www.sciencedirect.com/science/article/pii/S1755581720300110.

[39] M. Grieves, Digital twin: manufacturing excellence through virtual factory replication, 2014, pp. 1–7, White paper 1.

[40] F. Tao, H. Zhang, A. Liu, A.Y. Nee, Digital twin in industry: State-of-the-art, IEEE Trans. Ind. Inform. 15 (2018) 2405–2415, http://dx.doi.org/10.1109/TII.2018.2873186.

[41] D.J. Wagg, P. Gardner, R.J. Barthorpe, K. Worden, On key technologies for realising digital twins for structural dynamics applications, in: Model Validation and Uncertainty Quantification, vol. 3, Springer, 2020, pp. 267–272, http://dx.doi.org/10.1007/978-3-030-12075-7_30.

[42] S. Boschert, R. Rosen, Digital Twin—the Simulation Aspect, Springer International Publishing, Cham, 2016, http://dx.doi.org/10.1007/978-3-319-32156-1_5.

[43] L. Deren, S. Wenbo, S. Zhenfeng, Smart city based on digital twins, Comput. Urban Sci. 1 (2021) 1–11, http://dx.doi.org/10.1007/s43762-021-00005-y.

[44] K. Chaturvedi, T.H. Kolbe, Integrating dynamic data and sensors with semantic 3D city models in the context of smart cities, ISPRS Ann Photogramm. Remote. Sens. Spat. Inf. Sci. 4 (2016) 31–38, http://dx.doi.org/10.5194/isprs-annals-IV-2-W1-31-2016.

[45] R. Klar, N. Arvidsson, V. Angelakis, Digital twins' maturity: The need for interoperability, IEEE Syst. J. 18 (2024) 713–724, http://dx.doi.org/10.1109/JSYST.2023.3340422.

[46] M. Haraguchi, T. Funahashi, F. Biljecki, Assessing governance implications of city digital twin technology: A maturity model approach, Technol. Forecast. Soc. Change 204 (2024) 123409, http://dx.doi.org/10.1016/j.techfore.2024.123409, URL: https://www.sciencedirect.com/science/article/pii/S0040162524002051.

[47] A. Thaduri, Digital twin: Definitions, classification, and maturity, in: U. Kumar, R. Karim, D. Galar, R. Kour (Eds.), International Congress and Workshop on Industrial AI and eMaintenance 2023, Springer Nature Switzerland, Cham, 2024, pp. 585–599.

[48] Y. Liu, J. Feng, J. Lu, S. Zhou, A review of digital twin capabilities, technologies, and applications based on the maturity model, Adv. Eng. Inform. 62 (2024b) 102592, http://dx.doi.org/10.1016/j.aei.2024.102592, URL: https://www.sciencedirect.com/science/article/pii/S1474034624002404.

[49] S. Evans, C. Savian, A. Burns, C. Cooper, Digital twins for the built environment. An introduction to the opportunities, benefits, challenges and risks, 2019, white paper. Built Environment Institution of Engineering and Technology.

[50] Homa Masoumi, P.E. Sara Shirowzhan, C.J. Pettit, City digital twins: their maturity level and differentiation from 3D city models, Big Earth Data 7 (2023) 1–36, http://dx.doi.org/10.1080/20964471.2022.2160156.

[51] H.E. Pang, F. Biljecki, 3D building reconstruction from single street view images using deep learning, Int. J. Appl. Earth Obs. Geoinf. 112 (2022) 102859, http://dx.doi.org/10.1016/j.jag.2022.102859.

[52] B. Lei, R. Stouffs, F. Biljecki, Assessing and benchmarking 3D city models, Int. J. Geogr. Inf. Sci. 37 (2023b) 788–809, http://dx.doi.org/10.1080/13658816.2022.2140808.

[53] I. Paden, R. Peters, C. Garcia-Sanchez, H. Ledoux, Automatic high-detailed building reconstruction workflow for urban microscale simulations, Build. Environ. 265 (2024) 111978, http://dx.doi.org/10.1016/j.buildenv.2024.111978.

[54] J.M.D. Delgado, L. Oyedele, Digital twins for the built environment: learning from conceptual and process models in manufacturing, Adv. Eng. Inform. 49 (2021) 101332.

[55] J.M. Davila Delgado, L. Oyedele, Digital twins for the built environment: learning from conceptual and process models in manufacturing, Adv. Eng. Inform. 49 (2021) 101332, http://dx.doi.org/10.1016/j.aei.2021.101332, URL: https://www.sciencedirect.com/science/article/pii/S1474034621000859.

[56] C. Ellul, J. Stoter, B. Bucher, Location-enabled digital twins–understanding the role of NMCAS in a European context, ISPRS Ann. Photogramm. Remote. Sens. Spat. Inf. Sci. 10 (2022) 53–60.

[57] J. Stoter, K. Arroyo Ohori, F. Noardo, Digital twins: A comprehensive solution or hopeful vision? GIM Int.: Worldwide Mag. Geomatics (2021).

[58] C. Ellul, J. Stoter, B. Bucher, P.O. Olsson, R. Billen, B. DeLathouwer, C. Ellul, J. Stoter, B. Bucher, P. Olsson, et al., Towards national connected digital twins-a geospatial perspective, ISPRS Ann. Photogramm. Remote. Sens. Spat. Inf. Sci. 10 (2024) 147–154.

[59] M. Callcut, J.P. Cerceau Agliozzo, L. Varga, L. McMillan, Digital twins in civil infrastructure systems, Sustainability 13 (2021) 11549.

[60] J. Kirchherr, D. Reike, M. Hekkert, Conceptualizing the circular economy: An analysis of 114 definitions, Resour. Conserv. Recycl. 127 (2017) 221–232, http://dx.doi.org/10.1016/j.resconrec.2017.09.005.

[61] F. Reitsma, Revisiting the 'is giscience a science?' debate (or quite possibly scientific gerrymandering), Int. J. Geogr. Inf. Sci. 27 (2013) 211–221, http://dx.doi.org/10.1080/13658816.2012.674529.

[62] 23247-1, I., Automation systems and integration–digital twin framework for manufacturing–part 1: overview and general principles, 2020.

[63] E. Glaessgen, D. Stargel, The digital twin paradigm for future NASA and US air force vehicles, in: 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA, 2012, p. 1818.

[64] C. Semeraro, M. Lezoche, H. Panetto, M. Dassisti, Digital twin paradigm: A systematic literature review, Comput. Ind. 130 (2021) 103469, http://dx.doi.org/10.1016/j.compind.2021.103469, URL: https://www.sciencedirect.com/science/article/pii/S0166361521000762.

[65] E. Negri, L. Fumagalli, M. Macchi, A review of the roles of digital twin in cps-based production systems, Procedia Manuf. 11 (2017) 939–948, http://dx.doi.org/10.1016/j.promfg.2017.07.198, URL: https://www.sciencedirect.com/science/article/pii/S2351978917304067, 27th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017, 27-30 June 2017, Modena, Italy.

[66] E. VanDerHorn, S. Mahadevan, Digital twin: Generalization, characterization and implementation, Decis. Support. Syst. 145 (2021) 113524, http://dx.doi.org/10.1016/j.dss.2021.113524, URL: https://www.sciencedirect.com/science/article/pii/S0167923621000348.

[67] M. Dalibor, N. Jansen, B. Rumpe, D. Schmalzing, L. Wachtmeister, M. Wimmer, A. Wortmann, A cross-domain systematic mapping study on software engineering for digital twins, J. Syst. Softw. 193 (2022) 111361, http://dx.doi.org/10.1016/j.jss.2022.111361, URL: https://www.sciencedirect.com/science/article/pii/S0164121222000917.

[68] B. He, K.J. Bai, Digital twin-based sustainable intelligent manufacturing: A review, Adv. Manuf. 9 (2021) 1–21.

[69] C.H. dos Santos, J.A.B. Montevechi, J.A. de Queiroz, R. de Carvalho Miranda, F. Leal, Decision support in productive processes through des and abs in the digital twin era: a systematic literature review, Int. J. Prod. Res. 60 (2022) 2662–2681, http://dx.doi.org/10.1080/00207543.2021.1898691.

[70] M. Singh, R. Srivastava, E. Fuenmayor, V. Kuts, Y. Qiao, N. Murray, D. Devine, Applications of digital twin across industries: A review, Appl. Sci. 12 (2022) http://dx.doi.org/10.3390/app12115727, URL: https://www.mdpi.com/2076-3417/12/11/5727.

[71] I. Errandonea, S. Beltrán, S. Arrizabalaga, Digital twin for maintenance: A literature review, Comput. Ind. 123 (2020) 103316, http://dx.doi.org/10.1016/j.compind.2020.103316, URL: https://www.sciencedirect.com/science/article/pii/S0166361520305509.

[72] M. Soori, B. Arezoo, R. Dastres, Digital twin for smart manufacturing, a review, Sustain. Manuf. Serv. Econ. 2 (2023) 100017, http://dx.doi.org/10.1016/j.smse.2023.100017, URL: https://www.sciencedirect.com/science/article/pii/S2667344423000099.

[73] C. Coupry, S. Noblecourt, P. Richard, D. Baudry, D. Bigaud, Bim-based digital twin and xr devices to improve maintenance procedures in smart buildings: A literature review, Appl. Sci. 11 (2021) http://dx.doi.org/10.3390/app11156810, URL: https://www.mdpi.com/2076-3417/11/15/6810.

[74] S. Su, R. Zhong, Y. Jiang, Digital twin and its applications in the construction industry: A state-of-art systematic review, Digit. Twin 2 (2023) http://dx.doi.org/10.12688/digitaltwin.17664.2, version 2; peer review: 2 approved with reservations.

[75] D.G.J. Opoku, S. Perera, R. Osei-Kyei, M. Rashidi, T. Famakinwa, K. Bamdad, Drivers for digital twin adoption in the construction industry: A systematic literature review, Buildings 12 (2022) http://dx.doi.org/10.3390/buildings12020113, URL: https://www.mdpi.com/2075-5309/12/2/113.

[76] E. Shahat, C.T. Hyun, C. Yeom, City digital twin potentials: A review and research agenda, Sustainability 13 (2021) http://dx.doi.org/10.3390/su13063386, URL: https://www.mdpi.com/2071-1050/13/6/3386.

[77] A. Deprêtre, F. Jacquinod, A. Mielniczek, Exploring digital twin adaptation to the urban environment: comparison with cim to avoid silo-based approaches, ISPRS Ann. Photogramm. Remote. Sens. Spat. Inf. Sci. 4 (2022) 337–344.

[78] S. Shaharuddin, K.N. Abdul Maulud, S.A.F. Syed Abdu. Rahman, A.I. Ch. Ani, Digital twin for indoor disaster in smart city: A systematic review. The international archives of the photogrammetry, Remote. Sens. Spat. Inf. Sci. XLVI-4/W3-2021 (2022) 315–322, http://dx.doi.org/10.5194/isprs-archives-XLVI-4-W3-2021-315-2022, URL: https://isprs-archives.copernicus.org/articles/XLVI-4-W3-2021/315/2022/.

[79] C. Ghenai, L.A. Husein, M. Al Nahlawi, A.K. Hamid, M. Bettayeb, Recent trends of digital twin technologies in the energy sector: A comprehensive review, Sustain. Energy Technol. Assessments 54 (2022) 102837, http://dx.doi.org/10.1016/j.seta.2022.102837, URL: https://www.sciencedirect.com/science/article/pii/S2213138822008852.

[80] M.M.H. Sifat, S.M. Choudhury, S.K. Das, M.H. Ahamed, S. Muyeen, M.M. Hasan, M.F. Ali, Z. Tasneem, M.M. Islam, M.R. Islam, M.F.R. Badal, S.H. Abhi, S.K. Sarker, P. Das, Towards electric digital twin grid: Technology and framework review, Energy AI 11 (2023) 100213, http://dx.doi.org/10.1016/j.egyai.2022.100213, URL: https://www.sciencedirect.com/science/article/pii/S2666546822000593.

[81] J. Xia, G. Zou, Operation and maintenance optimization of offshore wind farms based on digital twin: A review, Ocean Eng. 268 (2023) 113322, http://dx.doi.org/10.1016/j.oceaneng.2022.113322, URL: https://www.sciencedirect.com/science/article/pii/S0029801822026051.

[82] A. Niaz, M.U. Shoukat, Y. Jia, S. Khan, F. Niaz, M.U. Raza, Autonomous driving test method based on digital twin: A survey, in: 2021 International Conference on Computing, Electronic and Electrical Engineering, ICE Cube, 2021, pp. 1–7, http://dx.doi.org/10.1109/ICECube53880.2021.9628341.

[83] F. Mauro, A. Kana, Digital twin for ship life-cycle: A critical systematic review, Ocean Eng. 269 (2023) 113479, http://dx.doi.org/10.1016/j.oceaneng.2022.113479, URL: https://www.sciencedirect.com/science/article/pii/S0029801822027627.

[84] M. Kosacka-Olejnik, M. Kostrzewski, M. Marczewska, B. Mrówczyńska, P. Pawlewski, How digital twin concept supports internal transport systems?—literature review, Energies 14 (2021) http://dx.doi.org/10.3390/en14164919, URL: https://www.mdpi.com/1996-1073/14/16/4919.

[85] Y. Lin, L. Chen, A. Ali, C. Nugent, C. Ian, R. Li, H. Ning, Human digital twin: A survey, 2022, arXiv preprint arXiv:2212.05937, arXiv:2212.05937.

[86] T. Nguyen, Q.H. Duong, T.V. Nguyen, Y. Zhu, L. Zhou, Knowledge mapping of digital twin and physical internet in supply chain management: A systematic literature review, Int. J. Prod. Econ. 244 (2022) 108381, http://dx.doi.org/10.1016/j.ijpe.2021.108381, URL: https://www.sciencedirect.com/science/article/pii/S0925527321003571.

[87] N.M. Sharef, N.A. Nasharuddin, R. Mohamed, N.W. Zamani, M.H. Osman, R. Yaakob, Applications of data analytics and machine learning for digital twin-based precision biodiversity: A review, in: 2022 International Conference on Advanced Creative Networks and Intelligent Systems, ICACNIS, 2022, pp. 1–7, http://dx.doi.org/10.1109/ICACNIS57039.2022.10055149.

[88] J. Wu, Y. Yang, X. Cheng, H. Zuo, Z. Cheng, The development of digital twin technology review, in: 2020 Chinese Automation Congress, CAC, 2020, pp. 4901–4906, http://dx.doi.org/10.1109/CAC51589.2020.9327756.

[89] G.P. Agnusdei, V. Elia, M.G. Gnoni, Is digital twin technology supporting safety management? a bibliometric and systematic review, Appl. Sci. 11 (2021) http://dx.doi.org/10.3390/app11062767, URL: https://www.mdpi.com/2076-3417/11/6/2767.

[90] D. Adamenko, S. Kunnen, R. Pluhnau, A. Loibl, A. Nagarajah, Review and comparison of the methods of designing the digital twin, Procedia CIRP 91 (2020) 27–32, http://dx.doi.org/10.1016/j.procir.2020.02.146, URL: https://www.sciencedirect.com/science/article/pii/S2212827120307800.

[91] K. Wang, Y. Wang, Y. Li, et al., A review of the technology standards for enabling digital twin, Digit. Twin 2 (2022a) 4, http://dx.doi.org/10.12688/digitaltwin.17549.2, URL: https://doi.org/10.12688/digitaltwin.17549.2.

[92] H. Hassani, X. Huang, S. MacFeely, Enabling digital twins to support the un SDGS, Big Data Cogn. Comput. 6 (2022) 115, http://dx.doi.org/10.3390/bdcc6040115.

[93] H. Wu, P. Ji, H. Ma, L. Xing, A comprehensive review of digital twin from the perspective of total process: Data, Model. Netw. Appl. Sensors 23 (2023) 8306, http://dx.doi.org/10.3390/s23198306.

[94] G. Mylonas, A. Kalogeras, G. Kalogeras, C. Anagnostopoulos, C. Alexakos, L. Muñoz, Digital twins from smart manufacturing to smart cities: A survey, IEEE Access 9 (2021) 143222–143249, http://dx.doi.org/10.1109/ACCESS.2021.3120843.

[95] Q. Qi, F. Tao, T. Hu, N. Anwer, A. Liu, Y. Wei, L. Wang, A.Y. Nee, Enabling technologies and tools for digital twin, J. Manuf. Syst. 58 (2021) 3–21.

[96] T. Deng, K. Zhang, Z.J.M. Shen, A systematic review of a digital twin city: A new pattern of urban governance toward smart cities, J. Manag. Sci. Eng. 6 (2021) 125–134.

[97] L. Raes, P. Michiels, T. Adolphi, C. Tampere, A. Dalianis, S. McAleer, P. Kogut, Duet: A framework for building interoperable and trusted digital twins of smart cities, IEEE Internet Comput. 26 (2021) 43–50.

[98] B. Li, W. Tan, A novel framework for integrating solar renewable source into smart cities through digital twin simulations, Sol. Energy 262 (2023) 111869.

[99] F.N. Abdeen, S. Shirowzhan, S.M. Sepasgozar, Citizen-centric digital twin development with machine learning and interfaces for maintaining urban infrastructure, Telemat. Inform. (2023) 102032.

[100] A. Zhang, J. Yang, F. Wang, Application and enabling digital twin technologies in the operation and maintenance stage of the AEC industry: A literature review, J. Build. Eng. 80 (2023) 107859, http://dx.doi.org/10.1016/j.jobe.2023.107859, URL: https://www.sciencedirect.com/science/article/pii/S2352710223020399.

[101] C. Liu, Y. Tian, Recognition of digital twin city from the perspective of complex system theory: Lessons from Chinese practice, J. Urban Manag. 12 (2023) 182–192.

[102] Digital Urban European Twins, D5.1 System Architecture & Implementation Plan, Technical Report, Digital Urban European Twins, 2020, URL: https://www.digitalurbantwins.com/_files/ugd/725ca8_d088894bf9ee466b9934463840d79689.pdf.

[103] A.L. Hananto, A. Tirta, S.G. Herawan, M. Idris, M.E.M. Soudagar, D.W. Djamari, I. Veza, Digital twin and 3D digital twin: Concepts, applications, and challenges in Industry 4.0 for digital twin, Comput. 13 (2024) 100.

[104] International Organization for Standardization, ISO/IEC 30173:2023 Digital Twin - Concepts and Terminology, Technical Report, International Organization for Standardization, Geneva, Switzerland, 2023, URL: https://www.iso.org/standard/81442.html.

[105] Y. Li, R. Krishnamurthy, S. Raghavan, S. Vaithyanathan, H. Jagadish, Regular expression learning for information extraction, in: Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing, 2008, pp. 21–30.

[106] Q. Le, T. Mikolov, Distributed representations of sentences and documents, in: International Conference on Machine Learning, PMLR., 2014, pp. 1188–1196.

[107] W.B. Cavnar, J.M. Trenkle, et al., N-gram-based text categorization, in: Proceedings of SDAIR-94, 3rd Annual Symposium on Document Analysis and Information Retrieval, Ann Arbor, Michigan, 1994, p. 14.

[108] S. Chaudhuri, K. Ganjam, V. Ganti, R. Motwani, Robust and efficient fuzzy match for online data cleaning, in: Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data, 2003, pp. 313–324.

[109] S.S. Gantayat, A. Misra, B. Panda, A study of incomplete data–a review, in: Proceedings of the International Conference on Frontiers of Intelligent Computing: Theory and Applications, FICTA 2013, Springer, 2014, pp. 401–408.

[110] G. Singh, S. Singh, Y. Daultani, M. Chouhan, Measuring the influence of digital twins on the sustainability of manufacturing supply chain: A mediating role of supply chain resilience and performance, Comput. Ind. Eng. 186 (2023) 109711, http://dx.doi.org/10.1016/j.cie.2023.109711, URL: https://www.sciencedirect.com/science/article/pii/S0360835223007350.

[111] R. Farhan Hussain, A. Mokhtari, A. Ghalambor, M. Amini Salehi, Chapter 1 - introduction to smart O&G industry: Overview of smart O&G industry, in: R. Farhan Hussain, A. Mokhtari, A. Ghalambor, M. Amini Salehi (Eds.), IoT for Smart Operations in the Oil and Gas Industry, Gulf Professional Publishing, 2023, pp. 1–17, http://dx.doi.org/10.1016/B978-0-32-391151-1.00010-1, URL: https://www.sciencedirect.com/science/article/pii/B9780323911511000101.

[112] H. Yeung, F. Kim, M. Donmez, J. Neira, Keyhole pores reduction in laser powder bed fusion additive manufacturing of nickel alloy 625, Int. J. Mach. Tools Manuf. 183 (2022) 103957, http://dx.doi.org/10.1016/j.ijmachtools.2022.103957, URL: https://www.sciencedirect.com/science/article/pii/S0890695522001080.

[113] A. El Saddik, Chapter 1 - introduction, in: A. El Saddik (Ed.), Digital Twin for Healthcare, Academic Press, 2023, pp. 1–13, http://dx.doi.org/10.1016/B978-0-32-399163-6.00006-8, URL: https://www.sciencedirect.com/science/article/pii/B9780323991636000068.

[114] H.H. Shahri, A.A. Barforush, A flexible fuzzy expert system for fuzzy duplicate elimination in data cleaning, in: International Conference on Database and Expert Systems Applications, Springer, 2004, pp. 161–170.

[115] M. Shafto, M. Conroy, R. Doyle, E. Glaessgen, C. Kemp, J. LeMoigne, L. Wang, Draft modeling, Simul. Inf. Technol. Process. Roadmap. Technol. Area 11 (2010) 1–32.

[116] K. Hribernik, T. Wuest, K.D. Thoben, Towards product avatars representing middle-of-life information for improving design, development and manufacturing processes, in: IFIP International Conference on Digital Product and Process Development Systems, Springer, 2013, pp. 85–96.

[117] R. Rosen, G. von Wichert, G. Lo, K.D. Bettenhausen, About the importance of autonomy and digital twins for the future of manufacturing, IFAC- Pap. 48 (2015) 567–572, http://dx.doi.org/10.1016/j.ifacol.2015.06.141, URL: https://www.sciencedirect.com/science/article/pii/S2405896315003808, 15th IFAC Symposium onInformation Control Problems inManufacturing.

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

[119] S. Haag, R. Anderl, Digital twin – proof of concept, Manuf. Lett. 15 (2018) 64–66, http://dx.doi.org/10.1016/j.mfglet.2018.02.006, URL: https://www.sciencedirect.com/science/article/pii/S2213846318300208, Industry 4.0 and Smart Manufacturing.

[120] H. Gohari, C. Berry, A. Barari, A digital twin for integrated inspection system in digital manufacturing, IFAC- Pap. 52 (2019) 182–187.

[121] M.N. Kamel Boulos, P. Zhang, Digital twins: from personalised medicine to precision public health, J. Pers. Med. 11 (2021) 745.

[122] N. Meier, R. Müller-Polyzou, L. Brach, A. Georgiadis, Digital twin support for laser-based assembly assistance, Procedia CIRP 99 (2021) 460–465.

[123] S.C.D. Pinto, Reality Anchor Methodology: Designing a Digital Twin Supporting Situation Awareness (Ph.D. thesis), université Paris-Saclay, 2022.

[124] T.T. Vo, M.K. Liu, M.Q. Tran, Harnessing attention mechanisms in a comprehensive deep learning approach for induction motor fault diagnosis using raw electrical signals, Eng. Appl. Artif. Intell. 129 (2024) 107643, http://dx.doi.org/10.1016/j.engappai.2023.107643, URL: https://www.sciencedirect.com/science/article/pii/S0952197623018274.

[125] N. Reimers, I. Gurevych, Sentence-bert: Sentence embeddings using siamese bert-networks, 2019, arXiv preprint arXiv:1908.10084, URL: https://arxiv.org/abs/1908.10084. presented at EMNLP 2019.

[126] Digital Twin Consortium, Digital twin consortium defines digital twin, 2020, URL: https://www.digitaltwinconsortium.org/2020/12/digital-twin-consortium-defines-digital-twin/. (Accessed: 07 November 2024).

[127] A. Trantas, R. Plug, P. Pileggi, E. Lazovik, Digital twin challenges in biodiversity modelling, Ecol. Inform. 78 (2023) 102357, http://dx.doi.org/10.1016/j.ecoinf.2023.102357, URL: https://www.sciencedirect.com/science/article/pii/S1574954123003862.

[128] L. Ehrlinger, J. Schrott, M. Melichar, N. Kirchmayr, W. Wöß, Data catalogs: A systematic literature review and guidelines to implementation, in: G. Kotsis, A.M. Tjoa, I. Khalil, B. Moser, A. Mashkoor, J. Sametinger, A. Fensel, J. Martinez-Gil, L. Fischer, G. Czech, F. Sobieczky, S. Khan (Eds.), Database and Expert Systems Applications - DEXA 2021 Workshops, Springer International Publishing, Cham, 2021, pp. 148–158.

[129] M. Abdelrahman, A. Chong, C. Miller, Personal thermal comfort models using digital twins: Preference prediction with bim-extracted spatial–temporal proximity data from build2vec, Build. Environ. 207 (2022) 108532, http://dx.doi.org/10.1016/j.buildenv.2021.108532, URL: https://www.sciencedirect.com/science/article/pii/S0360132321009240.

[130] M. Abdelrahman, C. Miller, Targeting occupant feedback using digital twins: Adaptive spatial–temporal thermal preference sampling to optimize personal comfort models, Build. Environ. 218 (2022) 109090, http://dx.doi.org/10.1016/j.buildenv.2022.109090, URL: https://www.sciencedirect.com/science/article/pii/S0360132322003274.

[131] J. Cho, C. Kim, K.J. Lim, J. Kim, B. Ji, J. Yeon, Web-based agricultural infrastructure digital twin system integrated with GIS and BIM concepts, Comput. Electron. Agric. 215 (2023) 108441, http://dx.doi.org/10.1016/j.compag.2023.108441.

[132] W. Wang, H. Guo, X. Li, S. Tang, J. Xia, Z. Lv, Deep learning for assessment of environmental satisfaction using BIM big data in energy efficient building digital twins, Sustain. Energy Technol. Assess. 50 (2022b) 101897, http://dx.doi.org/10.1016/j.seta.2021.101897.

[133] M. Bilal, L.O. Oyedele, J. Qadir, K. Munir, S.O. Ajayi, O.O. Akinade, H.A. Owolabi, H.A. Alaka, M. Pasha, Big data in the construction industry: A review of present status, opportunities, and future trends, Adv. Eng. Inform. 30 (2016) 500–521, http://dx.doi.org/10.1016/j.aei.2016.07.001, URL: https://www.sciencedirect.com/science/article/pii/S1474034616301938.

[134] C. Ramonell, R. Chacón, H. Posada, Knowledge graph-based data integration system for digital twins of built assets, Autom. Constr. 156 (2023) 105109, http://dx.doi.org/10.1016/j.autcon.2023.105109.

[135] Q. Lu, A.K. Parlikad, P. Woodall, G.D. Ranasinghe, X. Xie, Z. Liang, E. Konstantinou, J. Heaton, J. Schooling, Developing a digital twin at building and city levels: Case study of West Cambridge campus, J. Manag. Eng. 36 (2020) 05020004, http://dx.doi.org/10.1061/(ASCE)ME.1943-5479.0000763.

[136] C. Miller, M. Abdelrahman, A. Chong, F. Biljecki, M. Quintana, M. Frei, M. Chew, D. Wong, The Internet-of-Buildings (IoB)–digital twin convergence of wearable and IoT data with GIS/BIM, J. Phys.: Conf. Ser. IOP Publ. (2021) 012041.

[137] S. Agostinelli, A. Heydari, Chapter six - digital twin predictive maintenance strategy based on machine learning improving facility management in built environment, in: A.H. Elsheikh, M.E. Abd Elaziz (Eds.), Artificial Neural Networks for Renewable Energy Systems and Real-World Applications, Academic Press, 2022, pp. 149–158, http://dx.doi.org/10.1016/B978-0-12-820793-2.00007-0, URL: https://www.sciencedirect.com/science/article/pii/B9780128207932000070.

[138] A. Pal, J.J. Lin, S.H. Hsieh, M. Golparvar-Fard, Automated vision-based construction progress monitoring in built environment through digital twin, Dev. Built Environ. 16 (2023) 100247, http://dx.doi.org/10.1016/j.dibe.2023.100247.

[139] J. Xu, X. Shu, P. Qiao, S. Li, J. Xu, Developing a digital twin model for monitoring building structural health by combining a building information model and a real-scene 3D model, Measurement 217 (2023) 112955, http://dx.doi.org/10.1016/j.measurement.2023.112955.

[140] J. Kandt, M. Batty, Smart cities, big data and urban policy: Towards urban analytics for the long run, Cities 109 (2021) 102992, http://dx.doi.org/10.1016/j.cities.2020.102992, URL: https://www.sciencedirect.com/science/article/pii/S0264275120313408.

[141] B.N. Silva, M. Khan, K. Han, Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities, Sustain. Cities Soc. 38 (2018) 697–713, http://dx.doi.org/10.1016/j.scs.2018.01.053, URL: https://www.sciencedirect.com/science/article/pii/S2210670717311125.

[142] C. Lim, K.J. Kim, P.P. Maglio, Smart cities with big data: Reference models, challenges, considerations, Cities 82 (2018) 86–99, http://dx.doi.org/10.1016/j.cities.2018.04.011, URL: https://www.sciencedirect.com/science/article/pii/S0264275117308545.

[143] P. Liu, F. Biljecki, A review of spatially-explicit geoai applications in urban geography, Int. J. Appl. Earth Obs. Geoinf. 112 (2022) 102936, http://dx.doi.org/10.1016/j.jag.2022.102936.

[144] P. Liu, Z. Yan, F. Biljecki, Explainable spatially explicit geospatial artificial intelligence in urban analytics, Environ. Plan. B: Urban Anal. City Sci. 51 (2024a) 1104–1123, http://dx.doi.org/10.1177/23998083231204689.

[145] F. Biljecki, GeoAI for Urban Sensing, CRC Press, 2023, pp. 351–366, http://dx.doi.org/10.1201/9781003308423-17, chapter 17.

[146] S. Gao, Geospatial Artificial Intelligence (GeoAI), vol. 10, Oxford University Press, New York, 2021.

[147] M. Batty, Digital twins, Environ. Plan. B: Urban Anal. City Sci. 45 (2018) 817–820, http://dx.doi.org/10.1177/2399808318796416.

[148] C. Weil, S.E. Bibri, R. Longchamp, F. Golay, A. Alahi, Urban digital twin challenges: A systematic review and perspectives for sustainable smart cities, Sustain. Cities Soc. 99 (2023) 104862, http://dx.doi.org/10.1016/j.scs.2023.104862, URL: https://www.sciencedirect.com/science/article/pii/S2210670723004730.

[149] N. Korotkova, J. Benders, P. Mikalef, D. Cameron, Maneuvering between skepticism and optimism about hyped technologies: Building trust in digital twins, Inf. Manag. 60 (2023) 103787, http://dx.doi.org/10.1016/j.im.2023.103787, URL: https://www.sciencedirect.com/science/article/pii/S0378720623000356.

[150] N.G. Malek, M. Tayefeh, D. Bender, A. Barari, Live digital twin for smart maintenance in structural systems, IFAC- Pap. 54 (2021) 1047–1052, http://dx.doi.org/10.1016/j.ifacol.2021.08.124, URL: https://www.sciencedirect.com/science/article/pii/S2405896321008818, 17th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2021.

[151] M.A. Hariri-Ardebili, G. Mahdavi, L.K. Nuss, U. Lall, The role of artificial intelligence and digital technologies in dam engineering: Narrative review and outlook, Eng. Appl. Artif. Intell. 126 (2023) 106813, http://dx.doi.org/10.1016/j.engappai.2023.106813, URL: https://www.sciencedirect.com/science/article/pii/S0952197623009971.

[152] D. Lee, S. Lee, N. Masoud, M. Krishnan, V.C. Li, Digital twin-driven deep reinforcement learning for adaptive task allocation in robotic construction, Adv. Eng. Inform. 53 (2022) 101710, http://dx.doi.org/10.1016/j.aei.2022.101710, URL: https://www.sciencedirect.com/science/article/pii/S1474034622001689.

[153] C. Lo, C. Chen, R.Y. Zhong, A review of digital twin in product design and development, Adv. Eng. Inform. 48 (2021) 101297, http://dx.doi.org/10.1016/j.aei.2021.101297, URL: https://www.sciencedirect.com/science/article/pii/S1474034621000513.

[154] C. Verdouw, B. Tekinerdogan, A. Beulens, S. Wolfert, Digital twins in smart farming, Agricult. Sys. 189 (2021) 103046, http://dx.doi.org/10.1016/j.agsy.2020.103046, URL: https://www.sciencedirect.com/science/article/pii/S0308521X20309070.

[155] M. Konstantinidou, J.M.S. Grau, S. Sahala, N. Nesterova, J. de Kruijf, A methodology for using dynamic visualizations to enhance citizens engagement, Transp. Res. Procedia 72 (2023) 4026–4032, http://dx.doi.org/10.1016/j.trpro.2023.11.376, URL: https://www.sciencedirect.com/science/article/pii/S2352146523006737, . tRA Lisbon 2022 Conference Proceedings Transport Research Arena (TRA Lisbon 2022), 14th-17th November 2022, Lisboa, Portugal.