

## Digital twins, architectures, and elements in civil engineering – A multivocal literature review

Kay Smarsly, Muhammad E. Ahmad, Patricia Peralta, Thamer Al-Zuriqat, Heba Al-Nasser, Kosmas Dragos, and Carlos Chillón Geck

> Institute of Digital and Autonomous Construction, Hamburg University of Technology kay.smarsly@tuhh.de

#### Abstract

A plenitude of digital twin reviews have been published in recent years, most of which, however, solely review project-specific digital twin implementations or focus on specific application areas within civil engineering. Despite the widespread adoption of digital twin applications in recent years in civil engineering, there is a lack of agreement on a common definition of digital twins. A common definition and, furthermore, an understanding of digital twin architectures and the composition of the internal elements would advance digital twin implementations, enhancing the reliability and performance of digital twins. In this study, a multivocal review of digital twins in civil engineering is conducted, aiming to provide a scientific basis for reaching a consensus on the abstract architecture and the composition of the internal elements of digital twins. Methodologically, the review is conducted in a twofold manner, accounting for both peer-reviewed, indexed literature, as well as for material that may report on research and development projects and industrial applications not indexed in well-accepted citation databases. The main contribution of this study is a systematic overview of existing architectures of digital twins implemented in civil engineering, including the internal elements. Furthermore, a digital twin definition and a generic reference architecture are proposed. It is expected that the definition and the reference architecture will serve as blueprints for digital twin implementations in civil engineering, providing a common understanding and enhancing the reliability and performance of digital twins.

Keywords: Digital twins, digital twin elements, digital twin architectures, reference architecture, multivocal literature review.

## 1 Introduction

Digitalization of civil infrastructure has experienced remarkable advancements, driven by recent technological innovations [1]. Recent advancements include the digitalization of building operations using building information modeling [2], the integration of artificial intelligence (AI)-based smart sensors into structural health monitoring systems [3], and the introduction of mobile monitoring systems that build upon mobile robotic platforms [4]. In this context, digital twins play a pivotal role of increasing importance in propelling digitalization in civil engineering forward, by enabling engineers and analysts to create precise and dynamic digital representations of civil infrastructure that are updated using sensor data in real time [5]. Digital twins are the core components of modern engineering systems [6]. By seamlessly integrating real-time data from sensors with the digital twins, engineers and analysts can monitor structural behaviors as well as simulate various scenarios, predict potential issues, and plan maintenance activities proactively [7].

In recent years, a number of comprehensive digital twin literature reviews have been published, focusing on digital twin characteristics, applications, implications, concepts, and workflows, as well as summarizing recent advances and challenges. The interested reader is exemplarily referred to the most recent digital twin reviews published in [8-15], some of which explicitly putting emphasis on reviewing digital twin definitions [15]. Despite the efforts conducted in digital twin research and development, a clear definition of digital twins or a common understanding of architectures and internal elements of digital twins have yet to be proposed. A clear definition of digital twins and a common understanding of the architecture and the internal elements would advance (i) efficient communication between engineers (as the meaning and description of digital twins and digital twins would provide a foundation for the analysis of digital twins and would enable the recognition and mitigation of design issues), and (iii) reusability of digital twins (as a common understanding would facilitate interoperability). Furthermore, a generic digital twin reference architecture that defines, in an abstract manner, the arrangement of the internal elements may serve as a blueprint for digital twin implementations in civil engineering, which is expected to provide a common understanding and enhance the reliability and performance of digital twins, with significant benefits for researchers, practitioners, and policymakers.

To provide a definition as well as a panorama of the architectures and internal elements of digital twins, this paper presents a multivocal review of digital twins in civil engineering. From a methodological standpoint, the review follows a twofold approach that encompasses (i) peer-reviewed, indexed literature ("white literature") as well as (ii) non-indexed sources ("gray literature") that include industrial digital twin applications. Based on the review, a digital twins in civil engineering. The remainder of this paper is structured as follows: In the following section, the multivocal review is presented, including the research questions, the methodology, and the results. Then, the findings are summarized, followed by the recommendations that include the digital twin definition and the reference architecture. Finally, conclusions are drawn and potential future research to further advance digital twins in civil engineering is discussed.

# 2 A multivocal literature review of digital twins, architectures, and elements in civil engineering

This section presents the review methodology as well as the main results of the review. The main objectives of the review are to provide an understanding of digital twins within the context of civil engineering, focusing on definitions, architectures, and elements of digital twins. Specifically, a multivocal literature review (MLR) is conducted, which is favored over other review types as it amalgamates a systematic literature review (SLR), encompassing white literature, with a gray literature review (GLR). The GLR is important for reviewing digital twin architectures and elements that are featured in industrial digital twin developments not published in white literature. Clearly, gray literature on digital twins comprises millions of sources, which is why the definition of inclusion, exclusion, and quality assessment criteria is of utmost importance, as will be illuminated in the following subsections.

The review methodology, shown in Figure 1, involves three phases, (a) a planning phase, in which objectives of the MLR and procedures to achieve the objectives are defined, (b) an execution phase, in which studies are collected and data is extracted, and (c) a reporting phase, in which the data is analyzed. The phases of the review methodology, each comprising several steps, are presented in the following subsections.

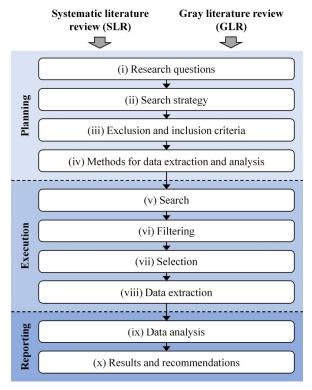


Fig. 1. Methodology of the multivocal literature review.

## 2.1 Planning phase

The planning phase includes four steps, (i) the definition of the research questions, (ii) the definition of a search strategy to retrieve studies from research and practice, (iii) the definition of inclusion and exclusion criteria to filter and select relevant studies, and (iv) the definition of methods for data extraction and data analysis. In the first step, the following research questions (RQs) are formulated in compliance with the objectives.

- RQ 1: What are the application areas of digital twins in engineering?
- RQ 2: What is the current state of the art in digital twinning in civil engineering?
- RQ 3: What definitions and characteristics of digital twins in civil engineering have been reported?
- RQ 4: What system architectures and elements have been reported?

In the second step of the planning phase, the search strategy is defined for the white literature (i.e., for the SLR) and for the gray literature (i.e., for the GLR). As for the SLR, the Scopus database serves as the data source because of its comprehensive coverage, quality assurance, and search tools. As for the GLR, to account for industry-related information not included in Scopus, the general-purpose open repository "Zenodo" [16], operated by the European Organization for Nuclear Research, is used as a data source. A relevancy ranking is then defined to bring the most relevant results to the top. In the third step, inclusion criteria (IC), exclusion criteria (EC), and quality assessment criteria (QAC) are defined, as listed in Tables 1, 2, and 3. The quality assessment criteria are scored with either QAC = 0 (a study does not satisfy the quality assessment criteria) or QAC = 1 (a study satisfies the quality assessment criteria). In the fourth step, the methods for data extraction and data analysis are defined. For example, title, document type, year, number of citations, discipline, and research focus are extracted and analyzed (to address RQ 1) or the definition of a digital twin is extracted and analyzed (to address RQ 3).

ID	Inclusion criteria
IC 1	The study meets the search terms "digital twin", "architecture", and "feature"
IC 2	The study has been published in English
IC 3	The study has been published in the period between 2013 and 2023
IC 4	The study has been published in conference proceedings or in academic journals
IC 5	The study is related to the areas of engineering, computer science, or mathematics

Table 2. Exclusion criteria for the multivocal literature review (SLR and GLR).

ID	Exclusion criteria
EC 1	The full text is not available
EC 2	The study mentions the term "digital twin", but does not elaborate on architectures

Table 3. Quality assessment criteria for the multivocal literature review (SLR and GLR).

ID	Quality assessment criteria
QAC 1	The objectives of the study are clearly stated
QAC 2	The limitations of the study are clearly stated
QAC 3	The methodology of the study is clearly stated
QAC 4	The outcomes of the study contribute to answering one or more research questions

#### 2.2 Execution phase

As shown in Figure 1, the execution phase encompasses four steps, (v) search, in which the studies are retrieved according to the research questions and the search strategy, (vi) filtering, in which the studies are filtered based on the inclusion and exclusion criteria, (vii) selection, in which the studies remaining after filtering are evaluated according to the quality assessment criteria so as to select relevant studies (hereinafter referred to as "selected studies"), and (viii) data extraction, in which data is extracted from the selected studies.

Regarding the SLR, the corresponding search strings are listed in Table 4. The search step, based on the search strings, initially yields 12,941 studies, which are narrowed down to 758 studies. In the filtering step, the 200 mostcited studies are identified, from which 70 meet the quality assessment criteria of the selection step. Additionally, 7 relevant studies are identified by snowballing, resulting in a total of 77 studies. After executing the data extraction step, in which the studies are systematically tabulated, the data extracted from the studies is forwarded to the reporting phase.

Regarding the GLR, the corresponding search string is listed in Table 5. In the search step, 830 studies are initially retrieved. By applying preliminary filters provided by Zenodo, also shown in Table 5, the result set is reduced to 137 studies. Upon conducting the actual filtering step, 49 studies are retained, 8 of which meet the quality assessment criteria of the selection step. Applying snowballing to the 8 studies yields 7 additional studies relevant to this review. Finally, the total of 15 studies are forwarded to the data extraction step, serving as input for the reporting phase.

**Table 4.** Search strings and results for the SLR within the multivocal literature review. For clarity, the restrictions, such as subject areas and language, are removed from the search strings displayed below.

No.	Search string	Results
1	("digital" AND ("twin" OR "shadow" OR "sibling"))	12,941
2	("digital" AND ("twin" OR "shadow" OR "sibling")) AND ("architecture" OR "framework" OR "platform")	4,822
3	("digital" AND ("twin" OR "shadow" OR "sibling")) AND ("architecture" OR "framework" OR "platform") AND ("feature" OR "service" OR "view")	1,435
4	("digital" AND ("twin" OR "shadow" OR "sibling")) AND ("architecture" OR "framework" OR "platform") AND ("feature" OR "service" OR "view") AND LIMIT-TO (EXACTKEYWORD ("digital twin" OR "internet of things" OR "iot" OR "industry 4.0" OR "cyber physical system" OR "architectural design" OR "architecture"))	758

Table 5. Search string and results for the GLR within the multivocal literature review.

Search string	Filters	Results
"digital twin" OR "digital twins"	Open access, project deliverable, report, technical note, and pdf	137

### 2.3 Reporting phase

The reporting phase includes two steps, (ix) data analysis, in which statistical data analysis is conducted, and (x) reporting of the results and recommendations, in which the conclusions regarding the research questions are drawn and recommendations are proposed. For clarity, in the following paragraphs, the main outcomes of both steps are presented together, and both the SLR and the GLR are included. First, an overview of the selected studies is provided. Then, the main findings in response to the research questions are summarized.

**Overview of the selected studies.** Interest in digital twins has exhibited a notable trend in recent years in research and practice that translates into the 92 selected studies in this review (comprising SLR and GLR), comprising a sample of relevant and most-cited studies in white and gray literature. Figure 2 shows the distribution of the selected studies. The statistical analysis of the data extracted from the selected studies has revealed an increase in the number of studies on digital twins since 2017/2018. The distribution of the 92 selected studies by document type is shown in Figure 3. The majority of the selected studies are journal papers (52%), followed by conference papers (20%), deliverables and reports (15%), reviews (12%), and white papers (1%).

0

White paper

2023

Gray literature Year Fig. 2. Distribution by year. Fig. 3. Distribution by document type.

Main findings. For clarity, the main findings of the MLR are summarized in response to the research questions.

2021

30

25

20

15 10

5

0

2013

2015

2017

2019

Count of the selected studies

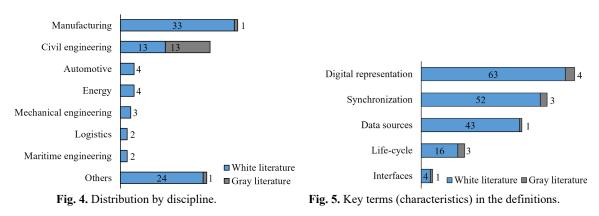
(White literature)

RO 1 ("What are the application areas of digital twins in engineering?") and RO 2 ("What is the current • state of the art in digital twinning in civil engineering?")

Regarding RQ 1, the majority of the selected studies are located within the application areas of manufacturing (34), followed by civil engineering (26), general topics related to digital twins (25), automotive (4), energy (4), mechanical engineering (3), logistics (2), and maritime engineering (2). Studies of multidisciplinary nature may involve two or more disciplines (Figure 4). Particularly in civil engineering (RQ 2), current trends in research have been reported that focus on the life-cycle stages of construction projects and on built environments, while trends in industry focus on providing services to satisfy practitioner needs, such as construction planning, inspections, multi-physics simulations, and monitoring. From the 26 studies classified into civil engineering (13 from the SLR and 13 from the GLR), most are related to the built environment (11), structural health monitoring (5), and facility management (4). Furthermore, it is observed that the studies in white literature mostly focus on modeling approaches to generate digital representations of physical assets, while the studies in gray literature primarily aim to advance "services" related to data management and integration, user interactions, and decision making.

RQ 3 ("What definitions and characteristics of digital twins in civil engineering have been reported?")

The research question RQ 3 concerns definitions of digital twins. A plethora of digital twin definitions and a broad wealth of reviews of digital twin definitions have been reported. For example, Barricelli et al. [8] presents a review of digital twin definitions, identifying key common terms used for defining digital twins, such as "virtual", "mirror", and "replica". In this paper, characteristics associated with key common terms are extracted to aid in the analysis of digital twin definitions. Specifically, from the selected studies, 67 definitions (63 from the SLR and 4 from the GLR) are available and are systematically analyzed by investigating the characteristics explicitly or implicitly mentioned in the definitions. As shown in Figure 5, "digital representation" is the most-reported characteristic, followed by "synchronization" and "data sources".



White literature

## • RQ 4 ("What system architectures and elements have been reported?")

The research question RQ 4 aims to investigate system architectures and internal elements of digital twins. In the selected studies, information on 78 architectures (70 from the SLR and 8 from the GLR) is provided. Essentially, the architectures are analyzed by investigating the layers (Figure 6) and the internal elements of the architectures (Figure 7), i.e. the functionalities embodied in the layers. As shown in Figure 6, six main layers are identified, (i) a modeling layer, hosting the digital representations, (ii) a service layer, providing services to interact with the digital representations, (iii) a perception layer, for data acquisition and device control, (iv) a data storage layer, for storing data, (v) a presentation layer, hosting user interfaces, and (vi) a transmission layer, for data transfer between the layers. Figure 7 shows the internal elements identified from the architectures. The most-reported internal elements are "devices" and "data storage", followed by elements devised to facilitate "services and tools", the "user interface", and "data analysis methods". In addition to responding to the research question, the modeling approaches pursued for describing the digital twin architectures are analyzed. As can be seen from Figure 8, the modeling approaches are categorized into three groups, (a) informal modeling techniques, such as box-and-line drawings, (b) semi-formal modeling languages, such as the Unified Modeling Language, and (c) formal languages, such as formal architecture description languages. As a result, none of the 78 architectures is modeled using formal modeling languages (74 informal, 4 semi-formal, 0 formal).

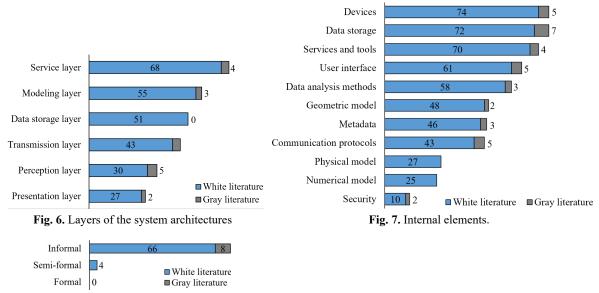


Fig. 8. Modeling approaches.

## **3** Results and discussion

In this section, the main findings are summarized and discussed, followed by the recommendations that include the digital twin definition and the generic reference architecture.

## 3.1 Summary and discussion of the results

From the results of the MLR, it is clear that digital twinning has been an emerging field in various domains and has been drawing increasing attention within civil engineering in research and practice. In addition to the high research interest, pronounced by the 77 studies yielded by the SLR, the GLR has underlined the growing engagement of the industry in developing the digital twin paradigm in civil engineering, as corroborated by the 15 studies that meet the quality assessment criteria of the MLR. Moreover, the increase in studies in the last years seems to be in tandem with the trends in digitalization in civil engineering, where the operation and maintenance of civil infrastructure have been aided by technologies and methodologies that have become state of the art in recent years, such as structural health monitoring and building information modeling. With digital twinning, heterogeneous data may be aggregated and integrated to describe built assets and the environment, thus meeting the demands and challenges of modern society.

Digital twins in research and practice share similarities across diverse definitions tailored for specific applications, as demonstrated by the common characteristics unveiled by the MLR, such as "digital representation" and "synchronization". However, digital twins in research primarily focus on aspects related to digital representations, while the industry gives higher importance to aspects related to synchronization, in an attempt to improve the fidelity of the digital twins and to support engineering services. Regarding digital twin architectures, most system architectures reported in the white and gray literature are inclined towards facilitating services, thus sharing similar structures, i.e. layers and internal elements. Remarkably, most digital twin systems in an intuitive manner, but fail to provide enough information for the technical implementation and analysis of the system architectures. None of the architectures analyzed in the MLR are modeled using formal modeling languages. Overall, it is expected that digital twinning will be widespread in future civil infrastructure projects, which highlights the importance of a commonly accepted definition and of a reference architecture of digital twins in civil engineering, which are recommended in the following subsections.

## 3.2 Recommendations

In view of the needs identified in the MLR, a general definition for digital twins is proposed that provides a basis for the understanding and the analysis of digital twins implemented in various civil engineering applications, including smart cities, smart buildings, and structural health monitoring. The following definition, recommended for digital twins in civil engineering and is intended to be kept rather general. The definition is distilled from the studies reviewed herein and, as such, reflects the perceptions interpreted from the studies:

## A "digital twin" is a digital representation of a real-world entity that dynamically mirrors and synchronizes with its real-world counterpart throughout either a part or the entirety of its life-cycle.

Similar to the digital twin definition, the reference architecture, recommended as follows, is synthesized from an analysis of the digital twin architectures presented in the studies. The reference architecture, shown in Figure 9, is organized into layers, adopting layered, service-oriented, and cloud-oriented design paradigms, i.e. similar functionalities are organized into layers. Flexibility in digital twin design is enabled by using the principle of "loose coupling", and digital twin scalability is provided through the ability to allocate cloud-based resources on the basis of business needs. More precisely, the reference architecture offers a clear distinction between the elements comprising the hardware ("data acquisition layer") used for procuring the data from the real world, the elements constituting the digital replica, i.e. the modeling/data storage/software elements ("platform layer"), and the visualization elements ("presentation layer") that belong to the digital world. The digital twin platform provides services that interact with the digital twins by querying, inserting, and processing data though the "transmission layer". The digital twins implemented in compliance with the reference architecture will be hosted in the modeling element, encompassing various digital representations, including building information models (BIM models), finite element (FE) models, sensor models, and performance models.

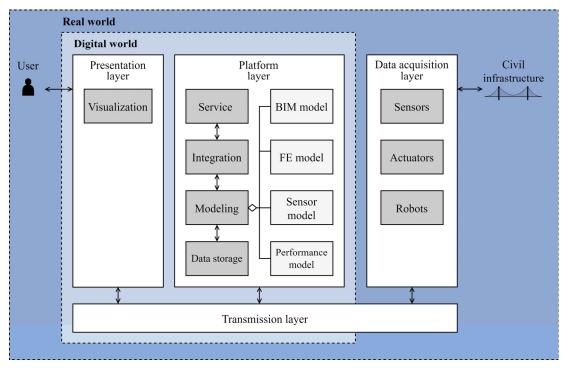


Fig. 9. Digital twin reference architecture.

## 4 Conclusions and future work

As the process of digitalization in civil engineering develops, the ubiquity of sensing devices and the subsequent availability of large amounts of data (e.g. structural response data and environmental data) is bound to further expedite the employment of digital twins in civil engineering applications. In this regard, and considering the increasing commitment of researchers and industrial stakeholders in advancing digital twins, the importance of gaining an overview of digital twinning in civil engineering, as well as of establishing widely accepted definitions for digital twins and a common understanding of internal elements and architectures, is pronounced. This paper has presented a multi-vocal literature review of digital twinning in civil engineering. The MLR has included both peerreviewed indexed literature sources, referred to as "white literature", and non-indexed sources, referred to as "gray literature". The review has been conducted in three phases, (a) the planning phase, in which the research questions, the inclusion/exclusion and the quality assessment criteria have been defined, (b) the execution phase, in which, upon retrieving and filtering literature, relevant studies have been selected and data thereof have been extracted, and (c) the reporting phase, in which the data has been statistically analyzed and reported, and recommendations have been formulated.

Out of more than 13,000 studies relevant to digital twins that have originally been retrieved in the MLR, a total of 92 studies have been selected as relevant, including 77 white-literature (SLR) studies and 15 gray-literature (GLR) studies, the latter primarily reporting on research and development projects and industrial applications. The review of the selected studies has provided a panorama of the research and practice in digital twinning in civil engineering. Moreover, the selected studies have been analyzed in terms of contents related to digital twin architectures. The analysis results have revealed traits shared among definitions of digital twins. Furthermore, the trend towards adopting formal schemas and cloud-computing paradigms is also reflected in the analysis results. In addition to the analysis, a definition for digital twins in civil engineering, on the basis of a common understanding of the internal elements, has been provided. Last, but not least, a reference architecture that is expected to serve as a blueprint for digital twin implementations in civil engineering, has been proposed.

The MLR, reported in this paper, has provided an overview of the state of the art in digital twinning in civil engineering. However, the purpose of the MLR has not been to conduct an exhaustive literature review of the field. In this direction, the outcome of the MLR could be extended to include further approaches, e.g. following the snowballing method, in an attempt to provide an exhaustive picture of digital twinning in civil engineering. Moreover, extending the literature review is expected to address the lack of formal system architectures for digital twins in civil engineering, aiming to enhance the reliability of digital twins and the trust of practitioners.

## Acknowledgments

This paper is the result of a collaborative effort of authors funded by different sources. The authors would like to express their sincere gratitude for the generous support received from the German Research Foundation (DFG) under grants SM 281/9-3, SM 281/17-1, SM 281/20-1, and SM 281/22-1, by the German Federal Ministry for Digital and Transport (BMDV) within the mFUND program under grant 19FS2013B, as well as by the German Federal Ministry of Health (BMG) through the Robert Koch Institute within the "AMELAG" project under grant 910003. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the sponsors.

## References

- 1. Opoku, D.-G. J., Perera, S., Osei-Kyei, R., & Rashidi, M., 2021. Digital twin application in the construction industry: A literature review. Journal of Building Engineering, 40(2021), 102726.
- Krämer, M. & Besenyői, Z., 2018. Towards Digitalization of Building Operations with BIM. IOP Conference Series: Materials Science and Engineering, 365(2), 022067.
- Luckey, D., Fritz, H., Legatiuk, D., Peralta, J., Walther, C., & Smarsly, K., 2022. Explainable artificial intelligence to advance structural health monitoring. In: Cury, A., Ribeiro, D., Ubertini, F., Todd, M. D. (eds.). Structural health monitoring based on data science techniques. Pp. 331-346. Cham, Switzerland: Springer.
- Smarsly, K., Dragos, K., Stührenberg, J., & Worm, M., 2023. Mobile structural health monitoring based on legged robots. Infrastructures, 8(9), 136.
- Richstein, R., Schmid, S., & Schröder, K.-U., 2022. Using SHM for the representation of structural components over their service life within digital twins. In: Proceedings of the 2022 European Workshop on Structural Health Monitoring. Palermo, Italy, July 4, 2022.
- Al-Nasser, H., Ahmad, M. E., Peralta, P., Chillón Geck, C., Al-Zuriqat, T., Dragos, K., & Smarsly, K., 2024. Digital twin architectures in civil engineering: A systematic literature review. In: Proceedings of the 15th Fachtagung Baustatik – Baupraxis. Hamburg, Germany, March 4, 2024.
- Taciroglu, E., Ebrahimian, H., & Ghahar, S. F., 2020. On the use of digital twins for structural health and performance monitoring and rapid post-event assessment. In: Proceedings of the 2020 PEER Annual Meeting – The Future of Performance-Based Natural Hazards Engineering. Berkeley, CA, USA, January 16, 2020.
- Barricelli, B. R., Casiraghi, E., & Fogli, D., 2019. A Survey on digital twin: Definitions, characteristics, applications, and design implications. IEEE Access, 7(2019), pp. 167653-167671.
- Pregnolato, M., Gunner, S., Voyagaki, E., De Risi, R., Carhart, N., Gavriel, G., Tully, P., Tryfonas, T., Macdonald, J., & Taylor, C., 2022. Towards civil engineering 4.0: Concept, workflow and application of digital twins for existing infrastructure. Automation in Construction, 141(2022), 104421.
- Schweiger, L. & Barth, L., 2023. Properties and characteristics of digital twins: Review of industrial definitions. 4(2023), 436.
- Sharma, A., Kosasih, E., Zhang, J., Brintrup, A., & Calinescu, A., 2022. Digital twins: State of the art theory and practice, challenges, and open research questions. Journal of Industrial Information Integration, 30(2022), 100383.
- Baghalzadeh Shishehgarkhaneh, M., Keivani, A., Moehler, R. C., Jelodari, N., & Roshdi Laleh, S., 2022. Internet of Things (IoT), building information modeling (BIM), and digital twin (DT) in construction industry: A review, bibliometric, and network analysis. Buildings, 12(10), 1503.
- da Silva Mendonça, R., de Oliveira Lins, S., de Bessa, I. V., de Carvalho Ayres, F. A. Jr., de Medeiros, R. L. P., & de Lucena, V. F. Jr., 2022. Digital twin applications: A survey of recent advances and challenges. Processes, 10(2022), 744.
- van Dinter, R., Tekinerdogan, B., & Catal, C., 2022. Predictive maintenance using digital twins: A systematic literature review. Information and Software Technology, 151(2022), 107008.
- 15. Semeraro, C., Lezoche, M., Panetto, H., & Dassisti, M., 2021. Digital twin paradigm: A systematic literature review. Computers in Industry, 130(2021), 103469.
- 16. European Organization for Nuclear Research (CERN), 2023. Zenodo. Accessed November 13, 2023, available at: https://zenodo.org/.