

A systematic review of digital twins: A business and management perspective

1. Introduction

Digital Twins (DTs) are attracting more attention as technology advances and they are being employed in a variety of scenarios. A DT serves as a virtual bridge between the physical world and the information world, which also fully uses models, data and simulations to provide real-time, efficient and intelligent services in multiple areas (Cimino et al., 2019; Zhang et al., 2023).

Companies have been challenged with digitalisation and rapid technological advances. In this context, technology implementation refers to the practical use of technology to enhance process performance. However, the implementation of digital twins brings about both organisational and individual changes. The success of implementing technologies is measured based on the time frame, budget and quality. Burton and Grange (2013) state that only effective use of technology can obtain maximum benefits. Organisations spend large amounts of money, so it is unsurprising that there are difficulties in realising the intended benefits only with financial benefits awareness. Breese (2015) indicates that the implementation is considered complete only when the desired benefits can be recognised and accompanied by short term gains. But the benefits can be achieved with appropriate management in a continuous process (Remenyi and Sherwood-Smith, 1998). Benefits cannot be captured when organisations do not know how digital twins can be utilised for this purpose, despite their technical success. Consequently, it is still quite challenging for an organisation to understand the benefits from the application of digital twins in terms of how to measure them, how they can affect corporate strategies, and business at the operational level. This review article aims to identify what managerial practices or improvements can increase the potential benefits from investing in digital twins and also the missing link between user response and the benefits from their use.

The rest of the paper is organised as follows: Section 2 represents the research approach and the method of selecting papers that focus on digital twins. It also presents the text-mining analysis of the content of the selected papers. Section 3 reviews the existing literature related to the main topics we obtained from the analysis. In section 4, we discuss the review conclusion and challenges for future research. Finally, the last section presents the limitations involved in the paper.

2. Methodology

The research adopts a systematic approach in the planning and literature review stage. The literature review process is a key tool to enable researchers to map and assess the existing intellectual territory, and to specify a research question to develop the existing body of knowledge further. To enhance the quality of the review and further

development of digital twins, the systematic review hopes to provide reliable evidence and relevant insights through theoretical synthesis into fields and sub-fields. However, traditional reviews lack a means of making sense of the collection of existing studies. Tranfield et al. (2003) emphasize the importance of a well-defined protocol, which includes clear research questions, search criteria, and data extraction procedures. In order to ensure that the findings are reached in a reliable and valid manner, the study followed a three-stage approach, as proposed by Tranfield, namely: the planning stage, conducting the review by analysing the papers and the report stage. The details are further discussed in this section.

2.1 Planning stage

Literature reviews vary depending on the types of research questions being investigated. So, the first step of the literature review should be a well-thought-out and planned outline. The planning stage aims to identify and refine the objectives of the research, which is focused on the organisational side of implementing DTs for business operation to gain benefits and sustainable operations. The clarified aim assists in searching for necessary types of information and sources, improving writing efficiency, and providing a well-structured article for readers.

2.2 Conducting stage

The main aim of the conducting stage is to filter the articles, evaluate and rate them to validate the relevance to the research objectives. The conducting stage consists of three steps: the paper selection, text mining and review.

2.2.1 Paper Selection

The Scopus database, one of the largest electronic databases of academic papers for research literature (Bar-Ilan, 2008), was used to find the literature for the review. The keywords for the search were 'digital' and 'twins' and the number of selected papers was 7353. Articles available in full text and in English were included in the initial search criteria. No limits were set on the publication year. The subject areas were limited to business. Consequently, the searching phrase is TITLE-ABS-KEY (digital AND twins) AND (LIMIT-TO (SUBJAREA , "BUSI")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (SRCTYPE , "j")) AND (LIMIT-TO (LANGUAGE , "English")). After applying the filters, the number of the results from Scopus was 296.

Preparing the analysis of the contents is the main aim of the text preprocessing. Vast amounts of raw data require new high-performance processing. However, low quality data will lead to low quality knowledge (Pyle, 1999). High-quality data is the foundation to assure the quality analysis results. Since data will be imperfect, containing inconsistencies and redundancies, it is not directly applicable for a data mining process. Thus, text preprocessing tasks, including data preparation and data reduction generally, convert the documents into intermediate forms, which should be suitable for different mining purposes (Zhang et al., 2015). Data preparation includes data transformation, missing value imputation, integration, cleaning and normalisation,

which aims to reduce the complexity of the data by feature selection, instance selection or by discretisation (García et al., 2016). Before proceeding to the next step, it is crucial to review the abstracts to ensure the quality and reliability of the research outcomes.

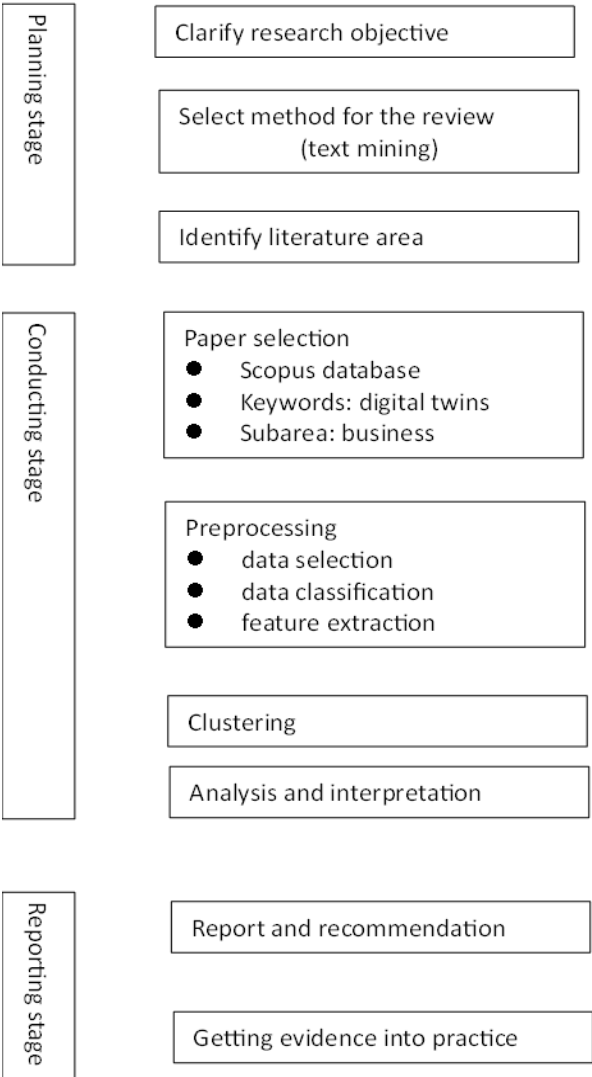


Table 1 Process of literature review (Tranfield et al. 2003).

2.2.2 Text mining

The text mining process aims to finish the information extraction and the result of analysis can usually be applied for specific purposes (Hearst, 2003). Quantitative content analysis has been applied for the content analysis. This analysis method can trace significance concepts. Critical aspects for the uptake of text mining technologies and tools in systematic reviewing are the existence of robust, scalable, efficient and rapidly responsive services for large collections and the necessity of consulting large-scale resources to support effective processing (Ananiadou et al. 2009). Generally, text mining is composed of three steps: text preprocessing, text mining operations, and clustering (Zhang et al., 2015).

For the text mining operations, the processed text was imported into WordStat (QDA miner add-on software) for quantitative full-text analysis. The reason for choosing

WordStat is because it enables the production of analysis results and QDA miner is a textual analysis software program. The technology is used to help maintain the scope of the review and filter down the resources to contain only the most relevant evidence. Word frequency analysis is known to be a text mining method to find the most frequently used words. Table 1 and table 2 list the most frequently occurring words and phrases in the reviewed literature.

Table 2: Term frequency

Term	FREQ	%CASES	TF IDF
DIGITALTWIN	787	80.34%	74.8
DIGITAL	303	43.39%	109.9
SYSTEM	223	37.97%	93.8
MANUFACTURING	219	29.15%	117.2
MANAGEMENT	178	28.47%	97.1
PROPOSED	176	35.93%	78.2
PHYSICAL	174	29.15%	93.1
PROCESS	170	30.17%	88.5
TIME	165	33.56%	78.2
INDUSTRY	160	28.81%	86.5
PRODUCTION	159	25.42%	94.6
SYSTEMS	157	29.83%	82.5
REAL	154	33.22%	73.7
SUPPLY	154	14.92%	127.3
SIMULATION	151	24.41%	92.5
INFORMATION	150	29.49%	79.5
TECHNOLOGIES	139	30.17%	72.3
TECHNOLOGY	135	28.14%	74.4
DEVELOPMENT	134	27.12%	75.9
DESIGN	131	24.07%	81

Table 3: Phrase frequency

Term	FREQ	%CASES	TF IDF
SUPPLY CHAIN	107	11.86%	99.1
REAL TIME	90	19.32%	64.3
DECISION MAKING	69	12.20%	63
ARTIFICIAL INTELLIGENCE	39	11.19%	37.1
SUPPLY CHAINS	33	6.44%	39.3
DIGITAL TWIN	32	10.85%	30.9
MACHINE LEARNING	32	8.47%	34.3
DIGITALTWIN TECHNOLOGY	30	8.14%	32.7
CYBER PHYSICAL	27	7.12%	31
DIGITAL TRANSFORMATION	27	6.44%	32.2
LIFE CYCLE	25	5.76%	31

MANUFACTURING SYSTEMS	21	4.41%	28.5
REAL WORLD	20	5.08%	25.9
INTELLIGENT MANUFACTURING	19	4.07%	26.4
BIM DIGITALTWIN	18	0.68%	39
DIGITAL TECHNOLOGIES	18	3.73%	25.7
PROPOSED FRAMEWORK	17	3.73%	24.3
SMART CITIES	17	3.39%	25
DIGITALTWIN SYSTEM	16	3.73%	22.9

For the clustering stage, fifteen research clusters were extracted from the WordStat in total, based on the keywords within each cluster, which would be further grouped. Table 3 illustrates the main keywords and phrases for each cluster, along with the grouping result. In order to conduct a comprehensive and a broader view of digital twins (thus, according to the main content of the discipline and keyword frequency), the clusters would be divided into two groups: “*Digital Twins: Adoption, Implementation, Benefits*” and the “*Applications of DTs*”. To facilitate a comprehensive understanding of the effective implementation process for benefit realisation, the clusters are systematically grouped into three main categories, aligned with the stages of adoption, implementation and benefit realisation. Firstly, the adoption group encompasses clusters that highlight factors influencing the decision-making process and the initial acceptance of digital twins. The implementation group includes the challenges and strategies associated with integrating digital twins into various systems. Finally, the benefit realisation group focuses on outcomes and measurable benefits derived from the successful deployment and utilisation of digital twins. As shown in table 4, the first eight items belong to the first group, while the rest of the items are classified into the second one. With the results from WordStat, a table was made to demonstrate the disciplines and keywords. According to the result from WordStat, it is necessary to determine the frequencies of key terms and phrases. The connections among the keywords based on the co-occurrence within the abstract were also analysed. For the research objective of the theme, particular attention has been paid to the first group, which will be more focused on the management side of the implementation, development, benefits, and adoption.

Table 4: Topic clusters

NO.	Dimension	Discipline	Keywords	FRE Q
1	Adoption	BIM; BUILDING	BIM; BUILDING; CYCLE; CONSTRUCTION; PROJECT; PROJECTS; MODELLING; ASSETS; TECHNICAL; IMPROVE;	309

			PLANNING; ENGINEERING; COMMUNICATION	
2	Adoption	REAL TIME	REAL; TIME; SYSTEM; SIMULATION; EXPERIMENTS; RMT; SCENARIOS	427
3	Implementatio n	ARTIFICIAL; INTELLIGENCE	ARTIFICIAL; INTELLIGENCE; LEARNING; AI; MACHINE; ML	193
4	Implementatio n	SCENARIO; ARCHITECTURE	SCENARIO; ARCHITECTURE; RESOURCES; CONDUCTED; ENABLED; DOMAIN; PLATFORM; BLOCKCHAIN	178
5	Implementatio n	PSS; OPTIMISATION	PSS; OPTIMISATION; MANUFACTURING; APPROACH;	170
6	Implementatio n & Benefits	TECHNOLOGIES SUCH INTERNETOFTHIN GS; BIG	INTERNETOFTHIN S; BIG; TECHNOLOGIES; COMPUTING; ADVANCED; CLOUD; SMART; INTELLIGENT; PARADIGM; INTELLIGENCE; CYBER; ARTIFICIAL	346
7	Benefits	DECISION MAKING	MAKING; DECISION; STRATEGIC; RISKS; RISK; DECISIONS	161
8	Benefits	CDC DAMPER; HYBRID	DAMPER; CDC; HYBRID; PARAMETERS; CONDITIONS; ACCURACY; LOW; DEVELOPING; ORDER; MODELING; CHARACTERISTICS	198
9	Application	CREATION;	CREATION;	231

		SERVICE	STUDIES; INNOVATION; SERVICES; SOLUTIONS; CUSTOMER; HEALTH; PUBLIC; PSS; SERVICE; INNOVATIVE	
10	Application	COVID-19 PANDEMIC; SC	PANDEMIC; COVID; SC; SUPPLY; RESILIENCE; CHAIN; DISRUPTION; CHAINS; RISK	207
11	Application	GREEN; TRANSITION	GREEN; TRANSITION; NATIONAL; DIGITALTWIN; TWIN; ECONOMY; STRATEGY; DIGITAL	490
12	Application	CITY; URBAN	CITY; URBAN; CITIES; SMART; ELDERLY; INFRASTRUCTURE	65
13	Application	ENTERPRISES; PRODUCTION	ENTERPRISES; PRODUCTION; ENTERPRISE; METHODS; CHARACTERISTICS; MINING	172
14	Application	SUPPLY CHAIN	SUPPLY; CHAIN; CHAINS; RISKS; FIRMS	57
15	Application	TRAFFIC; SCENARIOS	TRAFFIC; SCENARIOS; INFRASTRUCTURE; AIRPORT	10

2.2.3 Review analysis

Word frequency is the basic and most common method of data analysis. It examines the importance of words in a text or a set of texts by measuring how often certain words appear. In this part, the word frequency and the phrase frequency illustrate the most frequent words and phrases within the data from Scopus. “Digitaltwin”, “*Digital*” and “*System*” are the three most frequent words, and “*supply chain*”, “*real time*” and

“*decision making*” are the most frequent phrases.

Clustering subdivides the categories with different characteristics and then tries to find a mapping relationship between different categories (Sadeghi Moghadam et al., 2021). It also simplifies the complexity. Cluster extraction and keyword levelling revealed the relationship between similar terms (Ananiadou et al. 2009). The analysis identified fifteen clusters. Integrated with the research objective, the clusters are divided into two general groups, a management group and the application aspect. They include fifteen clusters in total and the review article will focus primarily on the management side.

2.3 Reporting stage

The final stage of the review was to report the descriptive statistics of the literature used in the review, and the findings of the analysis. In management research, the aim of conducting a review is to manage the diversity of knowledge for a specific academic inquiry and enable the researcher to develop the existing knowledge further. The systematic literature review utilising text mining has provided valuable insights into existing research into digital twins, as evidenced by the analysis of 296 articles extracted from Scopus. After the selection and preprocessing step, fifteen clusters were extracted from the software. After the analysis process, the most frequently-mentioned keywords were Digitaltwin (787), Digital (303) and system (223).

Clustering simplifies the complexity, and categorises keywords and content with different characteristics (Sadeghi Moghadam et al., 2021). Clustering works better with larger document collections as it can provide a more complete view. Clustering is available for providing a more accurate analysis result based on the frequency of terms. According to the WordStat statistics, fifteen clusters were obtained from the software. Integrated with the research objective, the clusters were divided into two groups and the review article will focus primarily on the management side.

3. Conclusion and future orientation

This study has conducted a systematic literature review that focuses on the adoption process of digital twins. The article took an in-depth look at the existing and expanding digital twin-related literature. It shows the motivations behind implementing smart digital twins, as well as a brief introduction to the application of digital twins. From a risk prevention perspective, the DT can put in place self-healing and self-configuring capabilities in order to minimise any disruption. However, the value of the DT concept and its implementation can vary between industries and different scenarios and use cases.

Increasing research and practical application provide a foundation for the maturity of digital twins. However, further systematic review research is expected to develop the

proposed model and provide a basis for the further DT implementation process with the continually updated database.

Ethical considerations should be further involved in the review research. The ethical implications of digital twin technologies encompass a series of issues, including data privacy, consent, and potential societal impacts. Examining how organisations handle sensitive data, ensure transparency, and incorporate ethical guidelines into their digital twin strategies is essential.

References

- Ananiadou, S., Rea, B., Okazaki, N., Procter, R., Thomas, J., 2009. Supporting systematic reviews using text mining. *Social science computer review* 27, 509–523.
- Bar-Ilan, J., 2008. Which h-index?—A comparison of WoS, Scopus and Google Scholar. *Scientometrics* 74, 257–271.
- Breese, R., Jenner, S., Serra, C.E.M., Thorp, J., 2015. Benefits management: Lost or found in translation. *International Journal of Project Management* 33, 1438–1451.
- Burton-Jones, A., Grange, C., 2013. From use to effective use: A representation theory perspective. *Information systems research* 24, 632–658.
- Cimino, C., Negri, E., Fumagalli, L., 2019. Review of digital twin applications in manufacturing. *Computers in industry* 113, 103130.
- García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J.M., Herrera, F., 2016. Big data preprocessing: methods and prospects. *Big Data Analytics* 1, 9. <https://doi.org/10.1186/s41044-016-0014-0>
- Hearst, M., 2003. What is text mining. *SIMS, UC Berkeley* 5.
- Pyle, D., 1999. *Data Preparation for Data Mining*. Morgan Kaufmann.
- Remenyi, D., Sherwood-Smith, M., 1998. Business benefits from information systems through an active benefits realisation programme. *International Journal of Project Management* 16, 81–98.
- Sadeghi Moghadam, M.R., Safari, H., Yousefi, N., 2021. Clustering quality management models and methods: systematic literature review and text-mining analysis approach. *Total Quality Management & Business Excellence* 32, 241–264.
- Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review - Tranfield - 2003 - *British Journal of Management* - Wiley Online Library [WWW Document], n.d. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-8551.00375> (accessed 1.9.24).
- Zhang, L., Guo, Y., Qian, W., Wang, W., Liu, D., Liu, S., 2023. Modelling and online training method for digital twin workshop. *International Journal of Production Research* 61, 3943–3962. <https://doi.org/10.1080/00207543.2022.2051088>
- Zhang, Y., Chen, M., Liu, L., 2015. A review on text mining. Presented at the 2015 6th

IEEE International Conference on Software Engineering and Service Science (ICSESS), IEEE, pp. 681–685.