Digital Twin Technologies, Architecture, and Applications: A Comprehensive Systematic Review and Bibliometric Analysis

Rosemary Ofosu[1], Amin Hosseinian-Far*[2], and Dilshad Sarwar[3]

[1, 2, 3]Department of Business Systems & Operations, University of Northampton, NN1 5PH, UK

[1]rosemary.ofosu93@gmail.com

[2, 3]{Amin,Hosseinianfar, Dilshad.Sarwar}@northampton.ac.uk

Abstract. Digital Twins, as a suite of technologies is progressively developing significant momentum in several fields of study. Various research works have been conducted outlining the concept, the underlying technologies, general and context-specific architectures, and applications. This study has been undertaken to identify relevant research areas, key authors, publishers, and geographical distribution of publications on digital twins through a systematic review and bibliometric analysis, to inform the trajectories of future research in the field. A keyword-based search for journals was first conducted in Web of Science Core Collection to obtain documents relevant for this study, and a systematic review was performed in accordance with the PRISMA guidelines. A bibliometric analysis was then performed on the extracted data using the VOSviewer software. The Tableau software and Microsoft Excel were also used to analyse and visualise some of insights derived from the analysis.

Keywords: Digital Twin, DT, Architecture, Reference Model, Digital Twin Application

1. Introduction

David Gelernter, an American computer scientist from the Yale University predicted digital twins back in the 1990s; He referred to the notion as “Mirror World” and described it as a technological voodoo figurine which would allow the world to be seen more profoundly, through the use of massive open software masterpieces (Gelernter, 1991). The concept, according to Grieves & Vickers (2017), gained recognition in the year 2002 when the University of Michigan presented the development of a Product Lifecycle Management (PLM) centre to industry. The conceptual idea for PLM had all the features of a digital twin: the tangible asset, the virtual asset, link for data flow from tangible asset to virtual asset, and vice versa. Digital twin can be described as the virtual delineation of a physical asset using data and stimulators for optimisation, real-time predictions, observing, and management, for superior decision making (Rasheed, et al., 2020). According to Maria (2020), the key components of digital twins are the model of the physical asset, the timeseries data captured with sensors from the asset, the unique identifiers that connect the physical asset and the virtual model, and its monitoring capability using cameras, sensors, etc. The implementation of digital twin facilitates cost reductions, improved product design (Crawford, 2021), and predictive and preventive maintenance (Daneshkhah, et al., 2017). In the bid to implement digital twin, various architecture and reference models have been proposed. As per Talkhestani, et al. (2019) and Aheleroff, et al. (2021), there has not been a clear holistic reference architecture for digital twin implementation: Different digital twin architectures have been proposed in different situations. Currently, digital twin is being applied in numerous sectors: Manufacturing, Healthcare, Production, Education, City Management, and many others.

There has been a number of bibliometric analyses in several aspects of digital twins: Ante (2021) performed a bibliometric analysis on digital twin technology in smart manufacturing and industry 4.0, Radanliev, et al. (2021) evaluated artificial intelligence and IoT cyber-physical systems in industry 4.0, and a bibliometric review was completed for digital twin enabled smart industrial systems by Ciano, et al. (2020), among others. This study however sought to comprehensively assess and analyse trends in overall digital twin research in terms of publication volumes, geographical dissemination, key authors, major journals, prominent publishers, countries and organisations, and research areas to determine the current and future trajectories for digital twin research. This was achieved by identifying and reviewing existing literature in relation to digital twins, its architecture, and applications, selecting keywords and conducting a systematic review, and adopting, and applying appropriate analysis tools for conducting a thorough bibliometric analysis, and discussing the findings. This helped to provide better insight on the state of existing literature in digital twins, and to inform future directions of research within the field.

2. Related Work

2.1. DT Definitions

Digital Twin (DT), which has been referred to as a fundamental enabler to digital transformation by Kritzinger, et al. (2018), has been defined by several researchers after its inception by Michael Grieves in 2002. In (Grieves & Vickers, 2017), digital twin was described as a holistic virtual view achieved by stripping information from a physical asset. Majumdar, et al. (2013) characterised digital twin as a fundamental paradigm that will include measurable data of material level attributes with high-level sensitivity. Wright & Davidson (2020) defined digital twin as a workable virtual model of a physical object. The concept was similarly described by Rosen, et al. (2015) as accurate models of the present state of a process and its conduct in collaborating with its ecosystem in the real world. Madni, et al. (2019) referred to the concept as a dynamic virtual model of a service, process, or system. Other researchers, including Barricelli, et al. (2019), Jones, et al. (2020) and Schleich, et al. (2017), in their definitions, related digital twin to computer-based models that mirror a physical system. This explains how the virtual model stimulates, emulates, and mirrors the physical asset using information assessed from the physical asset. On the other hand, Alam & Saddik (2017) and Schroeder, et al. (2016) referred to digital twin as "Part of a Cyber-Physical System" which may be defined as a group of physical units which have virtual components as their digital version, that work together with a virtual reality via a communication medium.

While one may be tempted to think that a digital twin is only a simulation or model, Grieves (2014), Kritzinger, et al. (2018) and Negri, et al. (2017) argue that it goes beyond that: A digital twin is an intelligent model which can evolve, and it follows the lifecycle of the physical asset. It allows predictions of future system failures and defects and facilitates simulation in order to test new configurations and facilitate predictive and preventive maintenance. The mirroring process is enabled by the harmonisation and continuous communication between the physical asset, its immediate environs, and its digital twin.

2.2. DT Characteristics

These authors in their definitions of digital twins had these characteristics in common: the physical object, the virtual model, and its interactions.

2.2.1. The Physical Object

This is the real-world artefact from which a digital replica is created. It may be an equipment, a component of an equipment, a process, or a living organism. Researchers often use specific terms such as 'product', 'component', 'system', etc to refer to these physical objects. As they are real-world objects, they are not usually characterised by the word 'physical'. In order to generalise and encompass all forms of physical objects, literature presents the use of some terms, including 'Physical Asset' (Huynh, et al., 2019; Paripooranan, et al., 2020), 'Physical Entity' (Barricelli, et al., 2019; Tao, et al., 2019), and 'Physical System' (Ketzler, et al., 2020). For the purpose of this study, all three terms will be used interchangeably. The physical entity has its own surrounding environment or real-world space within which it exists which includes all the parameters that may have an impact on the physical entity. This real-world space in which the physical entity exists has been named in literature as 'Physical Environment' (Leng, et al., 2019; Jones, et al., 2020). The physical environment includes the location, infrastructure, and technologies available, the time, and the status of the physical entity, among others.

2.2.2. The Virtual Entity

According to Bauer, et al. (2013), there are several kinds of virtual representations of physical assets: databases, 3D models, social media accounts, avatars, etc. However, the virtual entity is a controlled virtual representation of the physical asset that is precise on both a micro and macro level. Just as with the physical asset, the virtual entity is referred to by a number of terms for specificity; 'cyber', 'model', etc. For generalisation purposes Jones, et al. (2020) proposed the use of 'Virtual Entity'. Equally, the virtual entity has a surrounding environment which is a mirror the real-world environment. This has been popularly referred to as 'Virtual Environment' by researchers, including Grieves (2014) and Toivonen, et al. (2018). Lohtander, et al. (2018) described the virtual environment as parallel environment. This is because it precisely reflects the procedures and actions of the physical environment.

2.2.3. The Interaction Between the Physical Entity and the Virtual Entity

There is an endless connection between the physical entity and virtual entity in a digital twin. Barricelli, et al. (2019) explained that data is continually exchanged and revised as a result of real-time data uploads and big data analytics. Through this connection the state of the physical entity is conveyed to and mirrored by the virtual entity. Similarly, as explained by information flow and processes from the virtual entity is transmitted to and displayed by the physical entity. The physical entity and the virtual entity complement each other by facilitating data collection, storage and analysis from the entities and surrounding environment (Al-Ali, et al., 2020). The data, processes and information that flow between the physical entity and virtual entity are known as parameters and it is a two-way flow that can influence both entities.

2.3. DIGITAL TWIN ENABLING TECHNOLOGIES

There are new digital technologies springing up each day that support digital transformation. As part of this digital transformation, a technology that creates a virtual prototype of a physical entity has been introduced and is known as a digital twin technology (Lawton, 2021). Digital twins render unique visibility into physical systems and processes to be able to spot bottlenecks, be innovative and to restructure operations. As such, Aho (2020), termed digital twin as a facilitating technology for smart lifecycle management. According to Qi, et al. (2021), in the bid to create a digital twin of a physical entity, various digitalisation technologies have been employed. These digitalisation technologies include Internet of Things, Artificial Intelligence, Machine Learning, big data analytics, and cloud computing. The digitalisation technologies facilitate the converging of the physical and virtual entities of a digital twin.

2.3.1. Internet of Things

Internet of Things (IoT) is considered as one of the key enablers of digital twin. According to Nord, et al. (2019), IoT has no standard definition, although it has been defined by several researchers over the years. Lee & Lee (2015) defined it as a network of machines that can interact with each other. Correspondingly, Ornes (2016) characterised IoT as a connection of devices that keep growing and is able to capture and distribute data. Ben-Daya, et al. (2019) gave a more detailed definition by including features of connectivity, its nature that facilitates storage and sharing of data, and the communication among devices. They defined IoT as a network of real-world objects that are connected digitally to sense, observe, and collaborate in order to enable information sharing. These definitions refer to devices that are connected together and interact. The two components that these connected devices have in common are sensors to collect data and a means to analyse and communicate the collected data in real-time. This real-time analysis is key to digital twins as the sensors attached to the connected devices collect data and feed it to the digital twin instantaneously (Dave, 2020). Using this data, new concepts and logic are developed and tested on the digital twin.

2.3.2. Artificial Intelligence

Artificial Intelligence (AI) has been described by Amisha, et al. (2019) as the ability of computers and other technologies to mimic intelligent actions and critical thinking equivalent to a human. Despite the importance of intelligent systems, Teng & Gong (2018) argue that, without a learning ability it cannot be truly referred to as an intelligent system. The method through which a system is able to acquire knowledge on its own is known as Machine Learning (ML). Artificial intelligence and machine learning have become leading problem-solving methods, as drastic improvements in the capability and use of advanced analytical tools have changed the extraction of useful insights from big data. According to Dilmegani (2021), digital twins benefit from artificial intelligence and machine learning, since artificial intelligence and machine learning algorithms enable the development of some digital twins as well as the processing of big data gathered from digital twins. Machine learning frameworks facilitate the development of systems that can make independent decisions and make accurate predictions about future conditions using real-time data (Dohrmann, et al., 2019). Through machine learning, the processes become more intelligent, leading to more accurate management and analysis of complex data for better predictions.

2.3.3. Big Data Analytics

The large amounts of data generated from several different sources at high speed has necessitated big data analytics (Esmaeilbeigi, et al., 2020). Vassakis, et al. (2018) defined big data analytics as the collection, storage, assessment, and visualisation of large data sets to ascertain valuable insights to promote innovation, and transform businesses, as well as economies. Arunachalam, et al. (2018) explained that big data analytics is multidisciplinary and is characterised by its capability of managing data with four qualities: Volume (the size of the data), Velocity (the speed at which the data is produced), Variety (the format in which the data comes), and Veracity (the consistency of the data). Other researchers also characterise big data by the 3Vs (Volume, Velocity & Variety) (Zerhari, et al., 2015; Esmaeilbeigi, et al., 2020) or the 5Vs (Volume, Velocity, Variety, Veracity & Value) (Chen, et al., 2015; Polat, et al., 2019). According to Frankenfield (2020), most of the methods and processes involved in big data analytics are automated to evaluate data in order to identify critical patterns and metrics for better understanding. Big data analytics adopts different aspects of numerous scientific disciplines such as artificial intelligence, machine learning, statistics, system theory, and many others. As argued by IEEE Big Data (2018), big data analytics acts as an enabler for digital twins such that it allows creators of the digital twins to swiftly identify new development opportunities, make a diagnosis and rectify complications before they get out of control.

2.3.4. Cloud Computing

Cloud computing is one of the fastest emerging technologies in computing. National Institute of Standards and Technology (NIST) defined cloud computing as a model for facilitating accessible, on-demand system access to shared configurable computing resources such as servers, storage, applications, and networks with little or no service management interference. As explained by Hosseinian-Far, et al. (2018), the emergence of cloud computing is to address the need for businesses to collect and store huge amounts of quality data from numerous sources. Creating and managing a digital twin requires intensive computing and storage of data. According to Kumar, et al. (2018), cloud computing model consists of cloud provider, one who provides cloud services; a cloud consumer, one who gets the cloud service from the cloud provider; a cloud carrier, one who provides connectivity between the cloud provider and the cloud consumer; a cloud broker, one who interacts between the cloud provider and the cloud consumer to facilitate the business transaction; and the cloud auditor, one who independently assesses the cloud services, information systems, and performance and security operations. They explained further that the formalised service models are Infrastructure-as-a-Service

(IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS). Dohrmann, et al. (2019) expounded that, through the use of software-as-a-service (SaaS) solutions, which is a cloud-based software provision method for users, the cost of processing and storing the large amounts of data involved in digital twins continuous to decrease. It enables the developers of digital twins to acquire their needed computing resources as and when required at affordable costs. The computing power and resources necessary for real-time running simulations and forecasting of big data has been made widely and readily available by cloud computing.

2.4. DT Architecture

Generally, a digital twin architecture consists of a physical entity, its virtual entity, and a communication mechanism between the physical and virtual entities. As per Khan, et al. (2020), there is no single openly accepted architecture for digital twins. As such, several digital twin architectures have been proposed for different settings. The proposed digital twin architectures include 5C architecture, COGNITWIN, Intelligent digital twin, Six-Layer architecture, among others (Steindl, et al., 2020).

2.4.1. The 5-Layer Model of Cyber-Physical Systems

As the affordability and availability of computer networks, data acquisition systems and sensors are continually increasing, the use of high-tech methods has become a major force in several industries. These high-tech methods have led to the generation and use of large amounts of data, known as big data, accessed through these sensors and networked systems. According to Greer, et al. (2019) the technologies used to manage the networked systems between a physical entity and its computational abilities is known as a Cyber-Physical System (CPS). Lee, et al. (2015) proposed a 5-layer Cyber-Physical System known as the 5C architecture which gives a procedural guideline for creating and implementing a Cyber-Physical System for manufacturing applications. The five levels of the 5C architecture includes ‘Connection’, where effective sensors are selected, and precise and consistent data is acquired from manufacturing systems. Due to the availability of different types of data, a unified and tether-free technique for acquiring and transferring the data to the main server is useful. Also, it is crucial to select the most appropriate sensors for this level; ‘Conversions’, where important information is inferred from the raw data acquired. At this level, the architecture draws awareness to the health value and estimated useful life left for the machines involved through health management applications and prognostics; ‘Cyber’, which is the vital information centre such that it receives information from every connected device to form a network. Due to the massive information received on this level, the previous, current, and future performance of each device in the connection can be assessed to facilitate self-comparison among devices; ‘Cognition’, where infographics are used to transfer the vital information acquired to users. Here, decisions can be made on priority functions due to the availability of information on device statuses and comparative information; and ‘Configuration’, which is the response from the virtual space to the physical space and serves as an executive control to enable devices to self-configure and self-adapt. The corrective and preventive decisions undertaken in the Cognition Level is applied here (Figure 1).

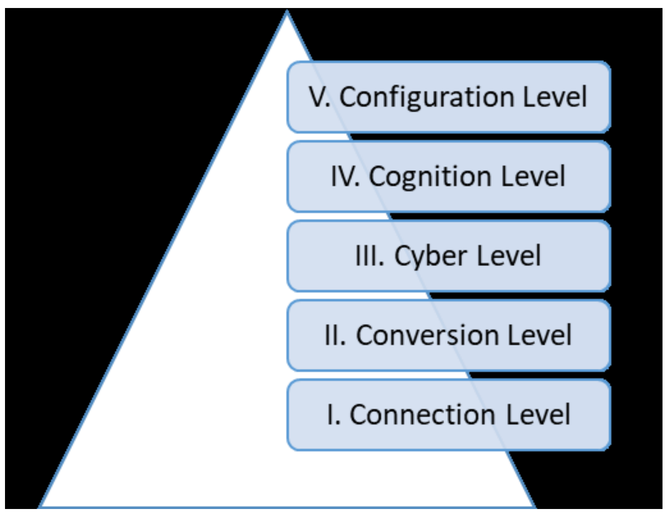

Figure 1: An image depicting the 5 Layer Cyber-Physical System/ 5C architecture adopted from (Lee, et al., 2015) and illustrated by author.

According to Lee, et al. (2015), employing a 5-Layer Cyber-Physical System in factories offer numerous benefits to production lines which consists of elements such as sensors, machinery, and production system. For the elements, the moment the sensory data from vital components has been transformed into useful information, a cyber-twin of the components will be liable for securing time machine records and integrating future measures to provide self-awareness and self-forecast. Subsequently, more complex machine data would be collected to the elements information to observe the status and create the cyber-twin of each device. These cyber-twins provide the extra self-assessment ability. Then with the production system, accumulated information from elements and device level information provide self-customisability and self-operability to the factory. At this point, the level of knowledge available guarantees a near zero production downtime and facilitates production and inventory planning for optimisation. Ahmadi, et al. (2017) asserts that the 5-Layer Cyber-Physical System facilitates engineering of better devices by leveraging performance data, remote device management, and operation optimisation, among other.

2.4.2. Six-Layer Digital Twin Architecture

A Six-Layer Digital Twin Architecture was created by Redelinghuys, et al. (2019) to facilitate interaction between the physical and digital entities, as well as the digital entity and the external world, targeted at circumstances where the products of several suppliers are utilised in the physical entity and the rest of the digital twin. According to Redelinghuys, et al. (2020), the Six-Layer digital twin architecture was inspired by the 5C architecture of Lee, et al. (2015). The first and second layers comprise of the physical entity. The first layer consists of physical devices such as sensors which exchange signals with the local controller. The local controller is located at the second layer which is the data source for the physical entity. The third layer provides a communication interface which is supplier neutral between the physical entity and the other layers of the architecture. Aside communicating with layer 4, it is able to directly log data to the cloud in layer 5 and at times layer six (Redelinghuys, et al., 2020). The fourth layer, also known as the IoT Gateway processes the data from the third layer to obtain useful insights for the upper layers. The fourth layer interfaces with the cloud data source, and the local data source. As such, adding a graphical user interface to this layer, where some of the major digital twin operations can be managed, is appropriate. The fifth layer is a cloud-based information repository for the physical and digital twin. As different stakeholders may have different information needs, numerous repositories are seen at this level. Holding the repositories in the cloud improves ease of access and use, and connectedness to the digital twin (Redelinghuys, et al., 2019). The sixth layer serves as a dashboard which connects the user to real-time historic information about the physical entity. It is equipped with emulation and simulation software such as Siemens Tecnomatix Plant Simulation which allows a user to interface with this layer.

2.4.3. COGNITWIN

Abburu, et al. (2020) presented an abstract architecture of the Cognitive Twin Toolbox (COGNITWIN) focusing particularly on the process industry. Three stages of twins were established: a Digital Twin, which makes use of only isolated prototypes of the physical system; a Hybrid Twin which has the ability to interrelate with its prototypes; and a Cognitive Twin which uses protracted prototypes that include proficient knowledge for problem-solving and to handle unfamiliar circumstances. The toolbox suggests five layers: Model Management Layer, Data Ingestion and Preparation Layer, Service Management Layer, Twin Management Layer, and a User Interaction Layer.

The required model types are almost parallel to the distinct semantic models in the reference framework for digital twin proposed by Josifovska, et al. (2019) and contains first-order theory models based on the underlying physics, empirical models, etc. The Service Management Layer is liable for managing services, like registering and planning. Two types of services are distinguished. Data-driven and model-based driven services are the two types of services recognised. The Twin Management Layer controls the composition of the digital twin, especially, the management problem as a result of changes in the performance of the physical system. The toolbox also presents a User Interaction Layer where clients can delve into the COGNITWIN.

2.5. DIGITAL TWIN APPLICATIONS

As a result of artificial intelligence, machine learning, Internet of Things, Big Data Analytics, and cloud computing working together, digital twins has become more detailed and predictive, thereby facilitating valuable applications. Digital twins have been applied in several sectors, ranging from healthcare to manufacturing. Some application cases emerging from literature are discussed as follows.

2.5.1. Manufacturing

According to Fuller, et al. (2020), the greatest reason for applying digital twins in the manufacturing sector is finding ways to track and monitor products while saving time and money. Other drivers for the application of digital twins in manufacturing include the need to gain competitive advantage, requirements for production flexibility on the market, the desire to follow a worldwide movement, the need to achieve process transparency, and safety concerns, among others (Neto, et al., 2020). As manufacturing processes are increasingly becoming digital, it is opening up opportunities for smart manufacturing. Qi & Tao (2018) explained that digital twins in manufacturing help to bridge the gap between the physical and virtual processes such that the use of IoT for the collection of real-time data in large volumes, based on cloud computing allows manufacturers to identify bottlenecks in their processes, trace to the root-cause and find the best possible solution. This ensures that manufacturing processes are efficient and more competitive. In accord, Xu, et al. (2019) also explained that digital twins provide a new concept for fault diagnosis such that issues in manufacturing processes which cannot be traced to its root-cause and assessed physical, can be evaluated on the virtual twin, with the appropriate what-if analysis to find the best solution.

2.5.2. Healthcare

Digital twins in healthcare is classified under digital health technology. Here, the physical entity may be living, may be in the form of wearable devices, software for diagnosis, medical devices, or drug development. Philips (2018) explored digital twins of a human heart. They focused on how clinicians could confidently assess disease states of the heart, determine treatment, and guide therapies enabled by anatomical intelligence. Liu, et al. (2019) developed a healthcare system referred to as cloud healthcare which is based on digital twin healthcare (CloudDIGITAL TWINH). It offers important insights into a design of a setup that consists of the patient, the physician, the digital twin, and the technical implementation of the digital twin prototype. With the digital twin healthcare potential solutions can be assessed in virtual environments, such as drug experimentations, preparations and simulations for surgeries, staff scheduling, etc. Also, Orcajo (2021) described wearables with sensors that feed real-time data to the cloud healthcare to enable patient monitoring and help develop model for the early detection of symptoms, diagnosis of diseases at its early stage and assess the effectiveness of treatment. Several drugs may also be tested on the digital twin patient to select the best drug for the situation, given the patient’s medical records and conditions.

2.5.3. Smart Cities

According to Kosowatz (2021), digital twins are being employed in the planning of cities to facilitate planning and prediction. The digital twin of a city is expected to reflect accurately and affect the laid down procedures used to operate and manage the city. As explained by Khajavi, et al. (2019), the virtual mirror of a city with all the constituents of the city represented on the virtual entity, provides an opportunity to improve operability and city planning. The major areas identified by Kumar, et al. (2020) as vital for developing smart cities are physical infrastructure, planning, information technology infrastructure, and smart solutions such as tourism services, tragedy management, etc. Smart cities are to improve residents' lives, promote security and environmental efficiency, through centrally regulated and supervised technical infrastructure (Niaros, et al., 2017). As asserted by Farsi, et al. (2020), digital twin technologies are important for city development and efficiency as it facilitates monitoring, the ability to clearly visualise, detecting and predicting concerns in real-time. Having the ability to view all nooks and crannies, and systems within a city virtually and applying what-if scenarios by leveraging on IoT technologies, artificial intelligence, and other technologies, to see what effects it will have on the city and its residents gives an opportunity to improve traffic flow, enhance energy efficiency, and improve security, among others. Some examples of smart cities given by Kosowatz (2020) include Singapore, which has integrated smart technologies into housing via a framework that takes into account, buildings, living, environment and planning to be able to analyse solar diffusion, wind flow, best sites for new constructions, etc; Dubai, which uses artificial intelligence to monitor bus drivers in order to reduce car accidents caused by exhaustion, and also having self-governing police stations where residents can make reports and pay for fines without dealing with a human being; and London, which desires to achieve a connected London by installing 5G cells 200 metres apart, using drones to identify places where cellular antennas can be installed, and fitting lampposts with sensors and charging ports for electric vehicles.

2.5.4. Education

Digital twin is becoming a new tool for education. According to Hinduja, et al. (2020), it allows rapid teaching and learning of new concepts. Using digital twins in education increases flexibility such that equipment that are too expensive for schools to afford and processes that are too slow to study physically may be accessed by just the click of a computer mouse. As explained by Hinduja, et al. (2020), digital twins in education encourages creativity and facilitates the simulation of complex experiments. In support, Sepasgozar (2020) also reasoned that digital twin in education improves creativity as there is the opportunity to have an active learning experience rather than a passive one. An example is students at Aarhus School of Marine and Technical Engineering using 3D replicas of automation systems to program automated production lines (Madni, 2019).

2.5.5. Transportation

As the populations increase, cities face the challenge of effective management of transport flows. There are usually control centres to ensure normal traffic flows on road networks and digitalising these centres is a key step to effective management. Rudskoy, et al. (2021) explained that digitalised control centres are achieved by implementing Intelligent Transport Systems. One of the major elements of the Intelligent Transport System is digital twins that make use of mathematical modelling methods to assess transport networks, identify issues and propose viable solutions. Access to the transport networks enables easy management of all aspects of the networks and the information related to it. According to Zhaohui, et al. (2021), digital twin in transportation ensures proper management of infrastructure, virtual assessments, and experiments to curb transportation issues.

3. Materials and Methods

This was quantitative research such that it employed quantitative methods for the study, that is, a systematic review and a bibliometric analysis. A systematic review was performed to obtain appropriate studies in digital twins for the research. A systematic review provides a summary of existing studies upon which informed judgements and recommendations can be made. A bibliometric analysis was also conducted to investigate the level of publications, the most researched subject areas, the most influential authors, and publishers, as well as the geographical distribution as used by Okumus, et al. (2019).

3.1. Identification of Sources and Data Collection

In identifying data sources, three major research databases were considered: Web of Science, Scopus, and Google Scholar. While some studies including Bramer, et al. (2017)

concluded that a single data source is not adequate for bibliometric analysis and systematic review, others such as Rice, et al. (2016) argued that a single database is sufficient as other databases have no impact on the results. A single database was selected based on the latter argument.

The data relevant to this study was retrieved from Web of Science Core Collection, an index of high-quality peer-reviewed publications currently managed by Clarivate Analytics. In order to identify all applicable publications, no lower limit was set, and the search was extended through to May 2021. The keyword string below was used to produce the preliminary database of publications in Web of Science.

“Digital Twin*” OR “DT” AND “Digital Twin* + Architecture” AND “Digital Twin* + Application” AND “Digital Twin* + Reference Model”

This keyword search yielded 2826 results. The quick filters on Web of Science were then used to filter the results on broad categories such as document type, the search results were limited to only journals; language, only the articles written in English were selected, and documents published later than May 2021 were excluded. This reduced the results to 1490 documents. The titles and abstracts of the publications were then manually assessed. The final dataset for this study after excluding additional 552 publications included a total of 938 publications. This search procedure was informed by PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) guidelines for performing systematic reviews (Figure 2).

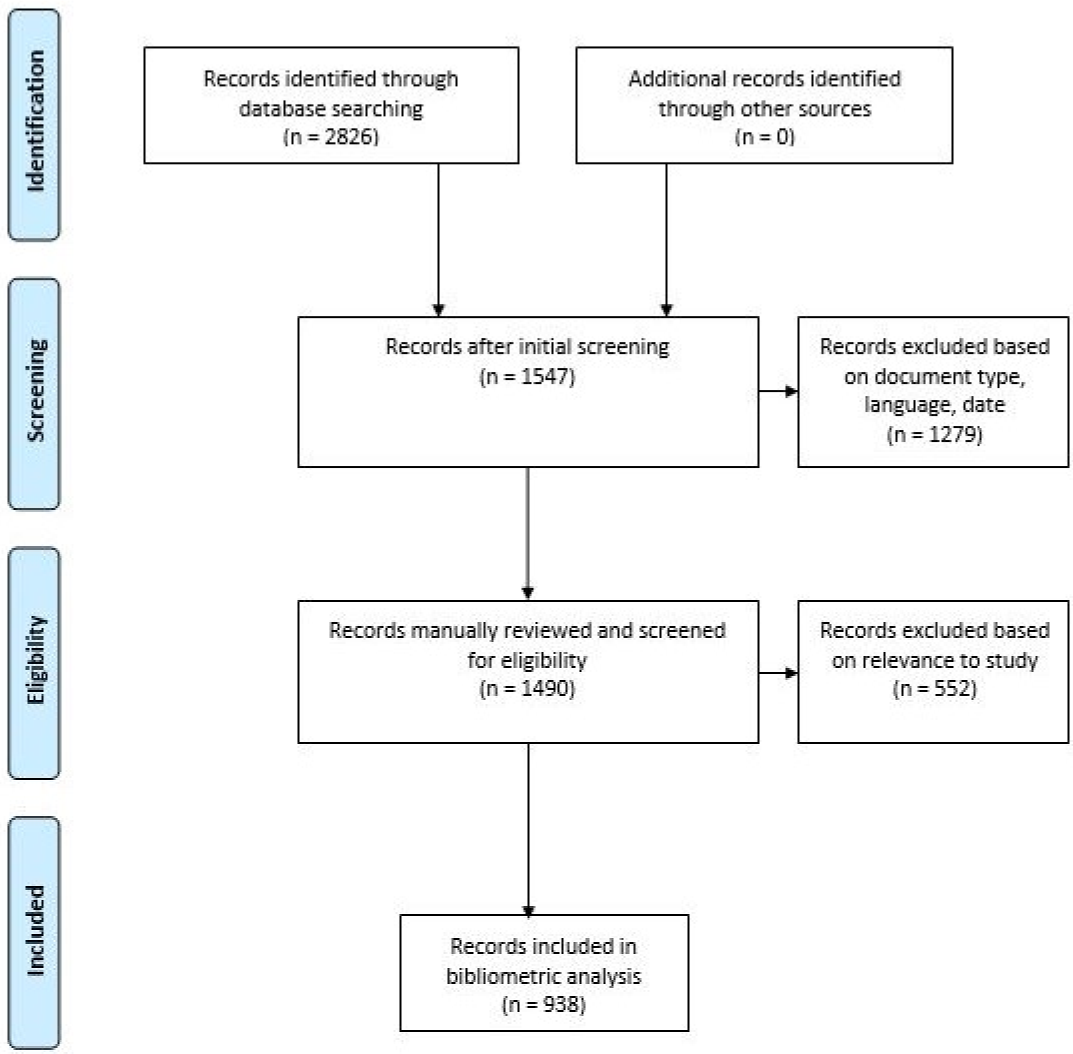

Figure 2: PRISMA flow diagram describing the collection of digital twin technology, architecture, and application documents from Web of Science (Moher, et al., 2009).

3.2. Data analysis methods

The data consisting of 938 documents was exported from Web of Science as both an Excel file and a Tab Delimited file. This is because the various analysis tools to be used could only work with one of these file types. The file consisted of meta data for each document, including name of authors, affiliated organisations and countries, title of article, source, abstract, publishers and publication locations, publication years, and other citation data. In order to use the VOSviewer software for the bibliometric analysis, a thesaurus file had to be created to filter the data, as used by Hallinger & Chatpinyakoop (2019). The thesaurus file had a 'label' column and a 'replace by' column, where same words expressed in a slightly different manner could be represented in singular form. For instance, some articles wrote ‘industry 4.0’ as ‘industry 4. 0’ and without the thesaurus file to correct this during the keyword analysis, VOSviewer was giving results for ‘industry 4.0’, ‘industry 4’ and ‘0’. Also, names such as ‘Tao, Fei’ represented as ‘Tao, F’ were being treated as different names. The quantitative analysis made use of descriptive methods, co-authorships, citation and co-citation analysis, and co-occurrence analysis. Descriptive methods such as the use of pie charts, bar charts, tree map charts and maps were used to present basic features such as publication growth trend, publication outlets and publication geographical distribution, and research areas. This was done with Excel and Tableau software programs.

VOSviewer, a software tool for structuring and visualising bibliometric systems such as journals and other publications was used to perform a bibliometric analysis. Jan van Eck & Waltman (2018) explain that VOSviewer has the ability to interpret and display large bibliometric maps in a comprehensible manner. Using the VOSviewer software, co-authorship analyses have been performed, a keyword co-occurrence has also been conducted to identified top concepts that have appeared across several of the research in relation to digital twins, and a citation and co-citation for authors and articles has been undertaken to identify key authors and key articles in digital twin research. According to Zupic & Čater (2015), the co-citation analysis complements the citation analysis as it captures a broader literature base by basing its analysis on the referenced list. As such, it is not uncommon for items to appear in a citation analysis list and not appear on a co-citation analysis list, especially if the citation analysis is based on just a single database.

4. Findings and Discussion

4.1. Findings

The results are presented in line with the deductive approach followed by Karakus, et al. (2019). It begins from more general findings and flows down to more specific results. It begins by giving details of the data such as the publications per year, research areas, and publications per publisher. From there, the findings on co-authorship follows as per the countries, organisations, and authors. The results on co-authorships are then presented according to countries, authors, and documents, after which the co-occurrences of author keywords are shown. Citation and co-citation of authors and documents are also presented. This makes it easier to understand as it gives broad knowledge about the data and findings before moving on to more specific outcomes.

4.1.1. Publication Trend

Out of the 938 documents retrieved from Web of Science, 1 was published as far back as 2004, 1 in 2014, 20 in 2017, 41 in 2018, 137 in 2019, 380 in 2020 and 310 from January 1 to May 31, 2021. 48 documents were early access journals released before May 31, 2021. Although the date for the search had no lower limits, there was no publication date older than 2004. As seen in Figure 3, there has been an upwards increasing trend in the publications on digital twin technology, its architecture, and applications because it recently started gaining thrust and the benefits derived from this concept has aroused the interest of industrial and research populations over the years (Pires, et al., 2019).

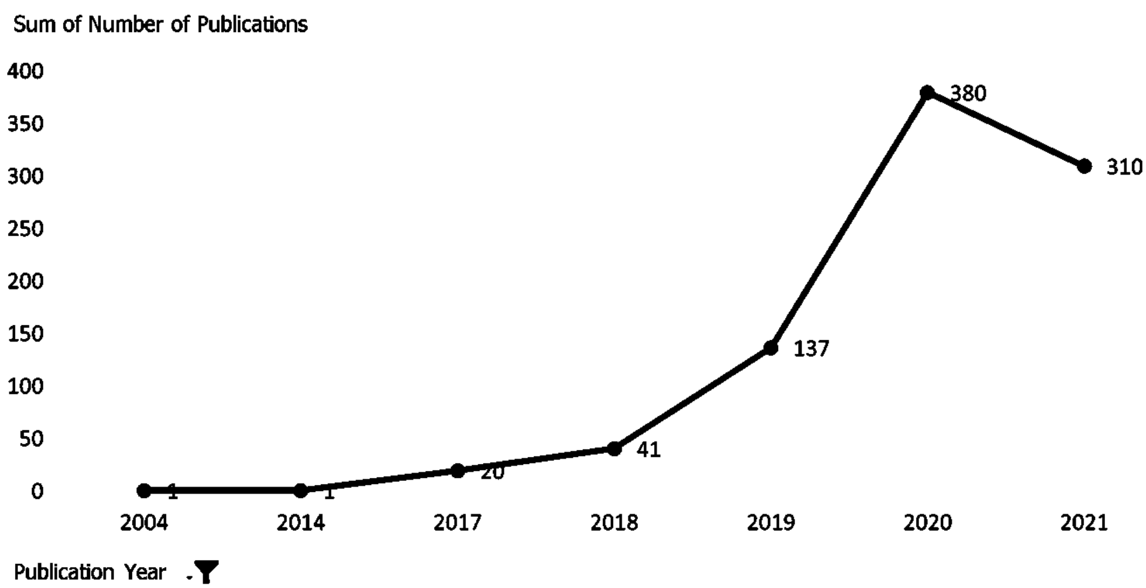

Figure 3: Growth trajectory of publications on digital twin technology, architecture, and applications (n=890).

This growth trend indicates that about 88% of the literature in digital twins have been produced in the past three years (2019 to 2021). This exponential growth is due to the realisation by industries and researchers the fascinating and substantial opportunities digital twin technologies offer (Ketzler, et al., 2020). These benefits of digital twins are projected to drive research over the years to come.

4.1.2. Publication Outlets

The 938 documents used for this study were published across 387 journals. The top 20 journals published 412 out of the 938, which makes 43.9% of the total number of journals. IEEE Access, which is a multidisciplinary peer-reviewed open access journal of the Institute of Electrical and Electronics Engineering (IEEE) was the most productive channel, publishing 63 journals in total. The next two were Applied Sciences (Basel, Switzerland) and Journal of Manufacturing Systems which published 46 and 44 documents respectively. The remaining journals published below 30 documents each out of the data used as shown in Figure 4.

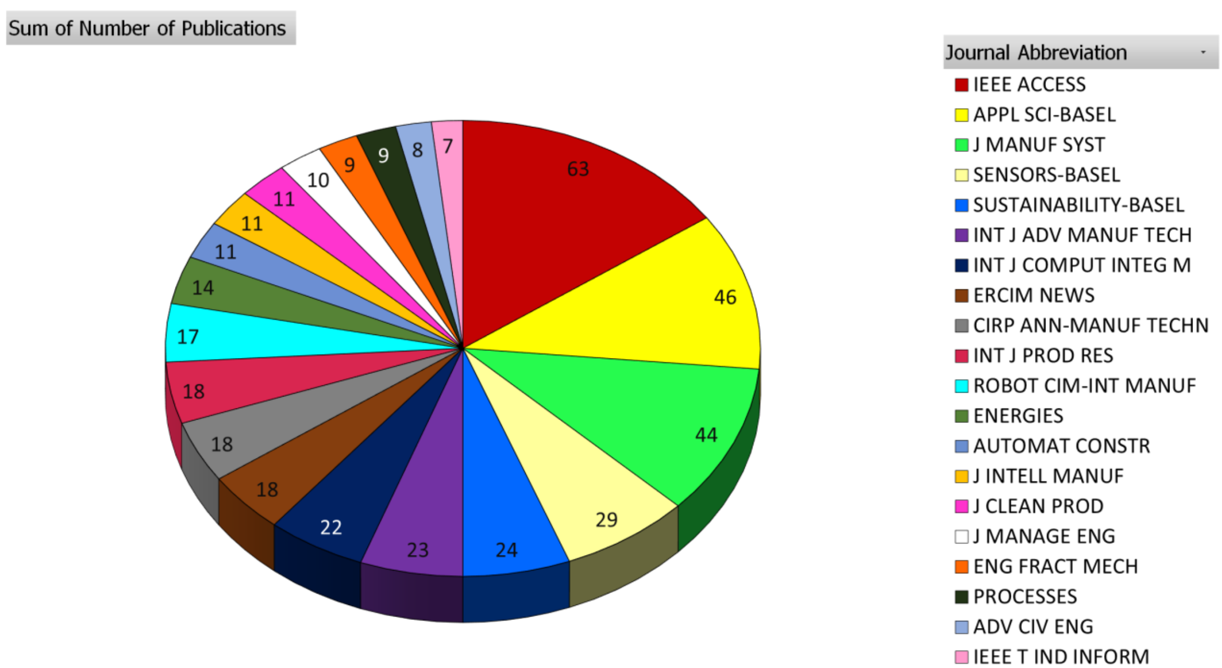

Figure 4: Top 20 journals in the publication of articles on digital twin technology, architecture, and applications ranked by number of articles in Web of Science until May 31, 2021.

The publishers of these documents were also evaluated and ranked using the number of publications. A total of 125 publishers were identified in relation to the 938 documents assessed. Some publishers had just 1 publication and it would be cumbersome to include all 125 publishers. As such, the top 20 publishers with the highest number of publications were selected and presented in Table 1 below. From the table, the publisher with the highest number of publications is MDPI (Multidisciplinary Digital Publishing Institute), a publisher of open access scientific journals, with 172 publications. The next was IEEE

(Institute of Electrical and Electronics Engineers) Publishing with 115 publications. Elsevier Science Ltd, Elsevier and Pergamon-Elsevier Science Ltd came next with 81, 79, and 60 publications respectively. Elsevier and Springer appeared a number of times in the top 20 but from different publication cities as shown in Table 1.

Table 1: Top 20 publishers in the publication of research on digital twin technology, architecture, and applications ranked by number of publications until May 31, 2021.

Publisher	Number of Publications	Publication City	Country
MDPI	172	Basel	Switzerland
IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC	115	Piscataway	USA
ELSEVIER SCI LTD	81	Oxford	England
ELSEVIER	79	Amsterdam	Netherlands
PERGAMON-ELSEVIER SCIENCE LTD	60	Oxford	England
TAYLOR & FRANCIS LTD	57	Abingdon	England
SPRINGER	25	New York	USA
SPRINGER LONDON LTD	25	London	England
SPRINGER HEIDELBERG	25	Heidelberg	Germany
WILEY	18	Hoboken	USA
EUROPEAN RESEARCH CONSORTIUM INFORMATICS & MATHEMATICS	18	Sophia Antipolis Cedex	France
EMERALD GROUP PUBLISHING LTD	14	Bingley	England
HINDAWI LTD	13	London	England
ASCE-AMER SOC CIVIL ENGINEERS	11	Reston	USA
ASME	10	New York	USA
SAGE PUBLICATIONS LTD	9	London	England
WILEY-HINDAWI	9	London	England
FRONTIERS MEDIA SA	7	Lausanne	Switzerland
WALTER DE GRUYTER GMBH	6	Berlin	Germany
ELSEVIER SCIENCE INC	6	New York	USA

The countries of the publishers were also assessed and represented in the form of a map in Figure 5 to show which countries where strongest in terms of publishing of articles on digital twin technology, its architecture, and applications. From the map, it is reflected that the United Kingdom had the most publications with 294 publications. United States of America (USA), Switzerland, and Netherlands followed with 221, 179, and 102 in that order.

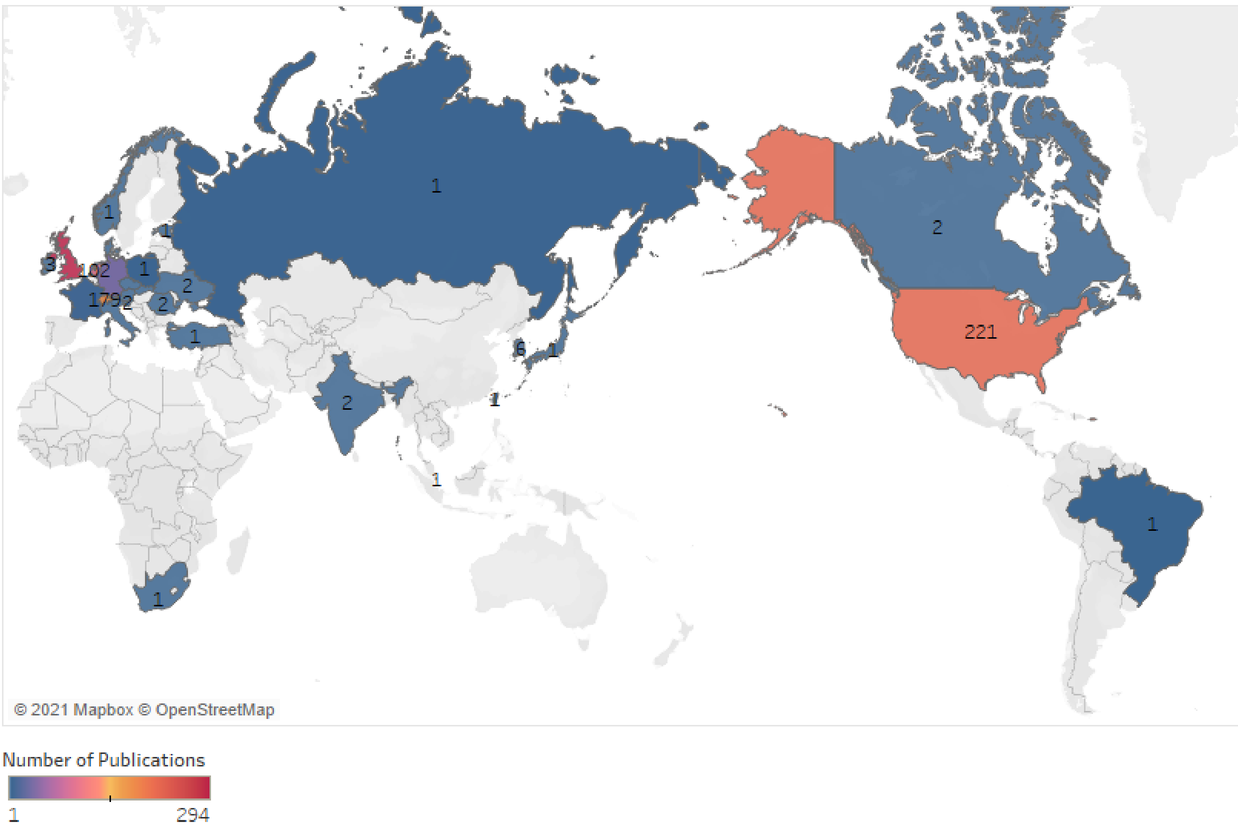

Figure 5: Geographical distribution of countries of publishers of research on digital twin technology, architecture, and applications.

4.1.3. Research Areas

The study was not limited to any specific research areas in order to ascertain meaningful representations in all fields in relation to digital twins as explained by Ross & Zaidi (2019). As such, 155 research areas were identified from the data extracted. The research area with the highest number of publications was Engineering, with 141 publications, which was almost twice as many publications as that of the next research area: Computer Science, Engineering & Telecommunications with 72 publications. The gap between the first two research areas was 69 publications, showing a clear advancement in research of digital twins in the field of engineering. Due to the large number of research areas identified, only the top 20 were represented in the Tree Chart below (Figure 6).

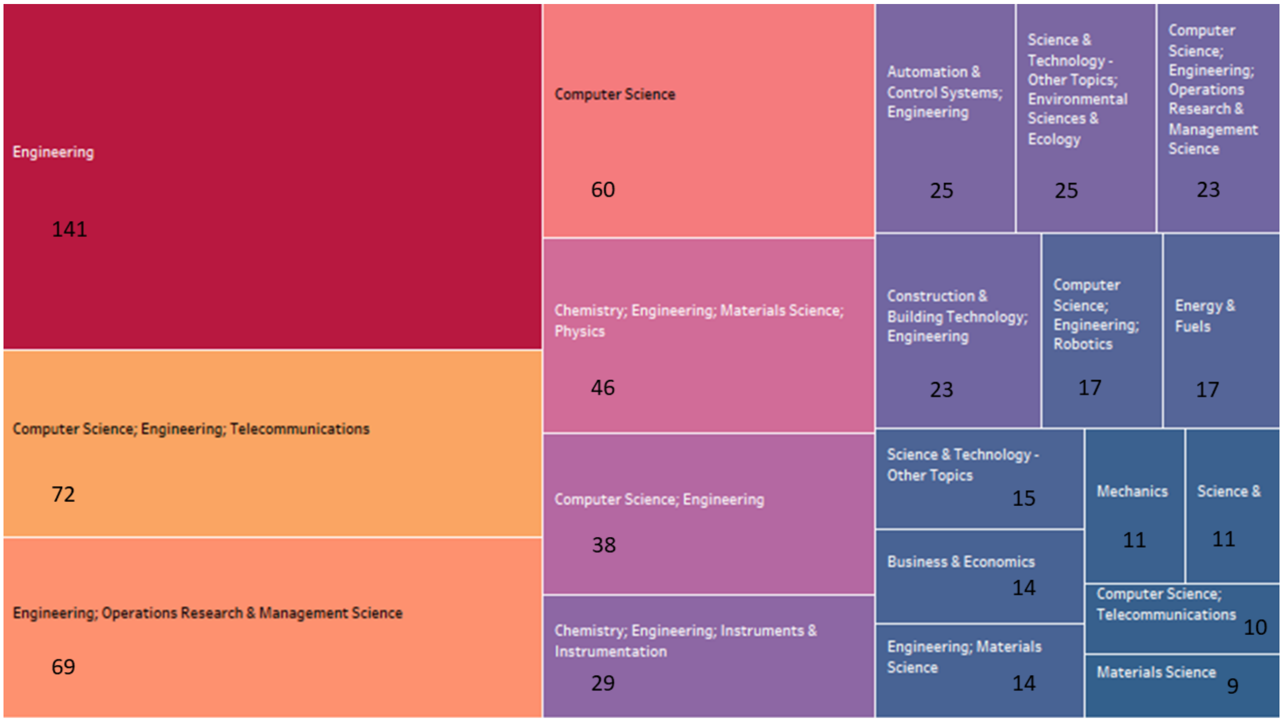

Figure 6: A Tree map chart presenting the top 20 research areas studied by the identified documents on digital twin research.

4.1.4. Co-authorship Analysis

According to Jalal (2019), co-authorship analysis is the assessment of the collaboration relationships between two or more research publications in a specified. It may be in terms of authors of the publications, its affiliated organisations, and/or countries.

The co-authorship of countries consists of countries, which are represented by nodes and links which connect the nodes in the form of co-authorships. There is a link between two countries if they have co-authored at least one document and the size of the nodes here are proportional to the total link strength of the country. For the purpose of this study, the minimum number of documents of a country was set to 5, and out of 70 countries associated with the publications, 39 met the threshold. The co-authorship of countries is presented with an overlay visualisation in Figure 7. Out of these 39 countries, the People’s Republic of China had the most co-authorships with 228 documents, 3495 citations, and a total link strength of 100. United States of America (USA) came in second with 139 documents, 1595 citations, and 89 as its total link strength. The details of the next 3 countries are specified in (X, Y, Z) format, made to represent number of documents, number of citations, and total link strength. The countries that followed are England with (81, 634, 63), Germany with (106, 930, 57), and France with (47, 620, 44). These were ranked as the top 5 countries for co-authorship based on their total link strength. The colour of the nodes in Figure 7 show how recent publications from the represented countries have been.

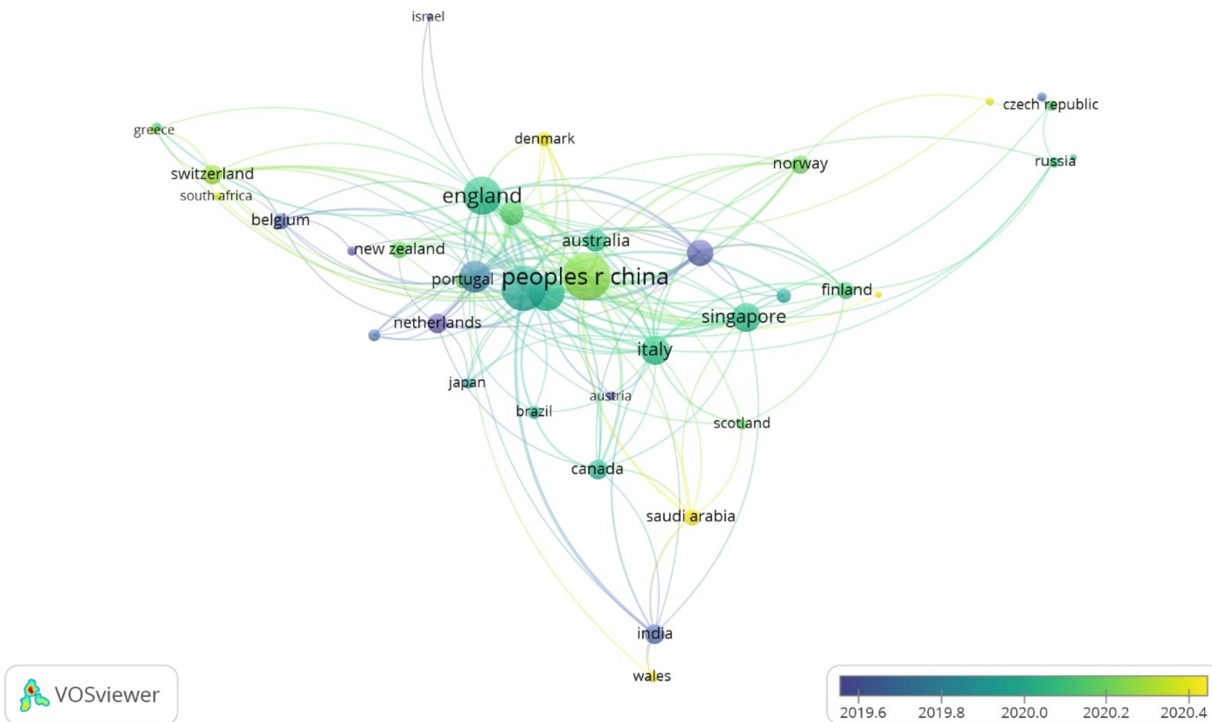

Figure 7: Country co-authorship overlay visualisation map of the literature on digital twin technology, architecture, and applications (n = 70 countries in the co-authorship network; threshold of 5 documents per country; display of 39 countries).

The co-authorship of organisations shows the relationship patterns between organisations related to co-authored documents. In this case, the organisations are represented by the nodes and the connections between the nodes are shown by the links. Setting the minimum number of documents of an organisation to 10, 18 organisations out of the 1159 organisations associated with the publications in the data met the threshold. Out of these 18 organisations, the largest set of connected nodes consisted of 11 organisations as shown in the network visualisation in Figure 8. The total number of documents, citations and total link strengths for each organisation was calculated. The organisation with the highest total strength was University of Hong Kong with 15 documents, 92 citations, and 12 for total link strength. This was followed by Beihang University with 23 documents, 1364 citations and 8 for total link strength. The other organisations are presented in the format (X, Y, Z), representing number of documents, citations, and total link strength respectively. Jinan University followed with (12, 27, 8), National University of Singapore with (14, 738, 8), University of Auckland with (11, 324, 4), Nanyang Technological University with (12, 217, 3), Guangdong University of Technology with (11, 530, 2), Zhejiang University with (10, 46, 2), Beijing Institute of Technology with (10, 167, 1), Sungkyunkwan University with (11, 107, 1), and then University of Cambridge with (17, 230, 1). The organisations stated were the top 11 organisations in terms of total link strength which constituted the largest set of connected nodes.

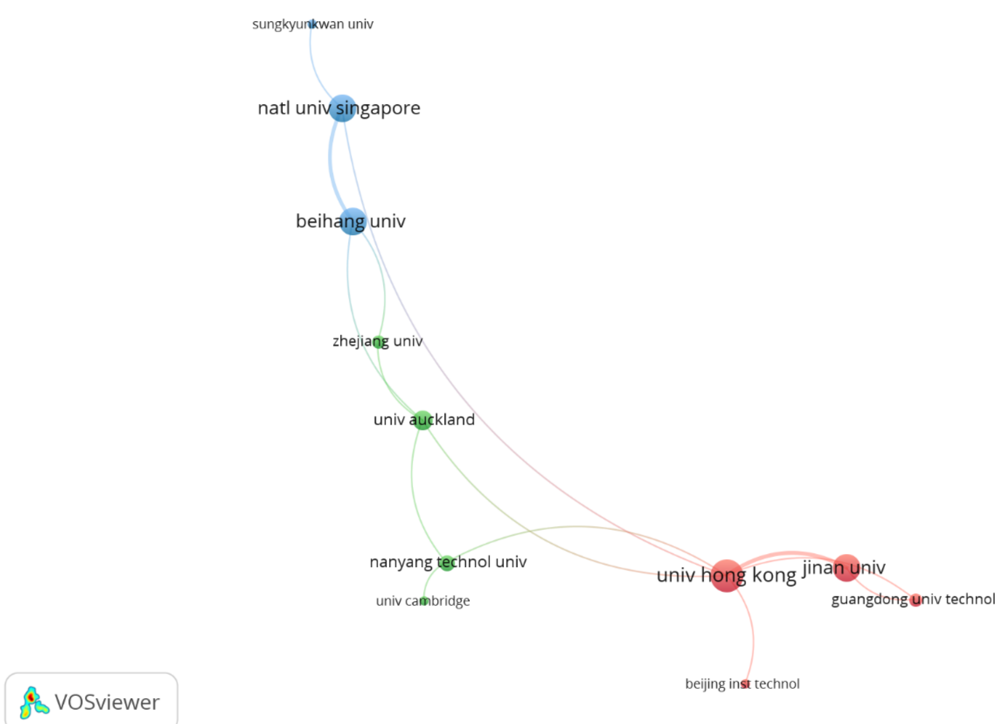

Figure 8: Organisation co-authorship network visualisation map of the literature on digital twin technology, architecture, and applications (n = 1159 organisations in the co-authorship network; threshold of 10 documents per organisation; display of 11 organisati

Author co-authorships of the publications are presented with a density visualisation in Figure 9. For inclusion purposes, the minimum number of documents of an author was set to 5. Of the 3410 authors identified in the dataset, 25 met the threshold. In Figure 9, the density visualisation was weighted by the total link strength and the portions turning yellow signify a larger total link strength. Using the total link strengths, the top 10 authors with strongest co-authorships are presented below. The author with the strongest co-authorship was Qiang Liu with 8 documents, 475 citations and total link strength of 29. Xin Chen and Jiewu Leng followed, each with 7 documents, 461 citations, and a total link strength of 29. As shown in Figure 9, these 3 authors have the brightest and biggest portion of yellow in the density visualisation. The specifics of the remaining of the top authors in co-authorships are presented in the format (X, Y, Z). Ding Zhang followed with (5, 331, 23), then Douxi Yan with (5, 240, 22), Fei Tao with (11, 1266, 12), Rikard Soderberg with (8, 197, 12), Kristina Warmefjord with (7, 187, 11), A.Y.C. Nee with (6, 653, 10), and Lars Lindkvist with (5, 181, 9).

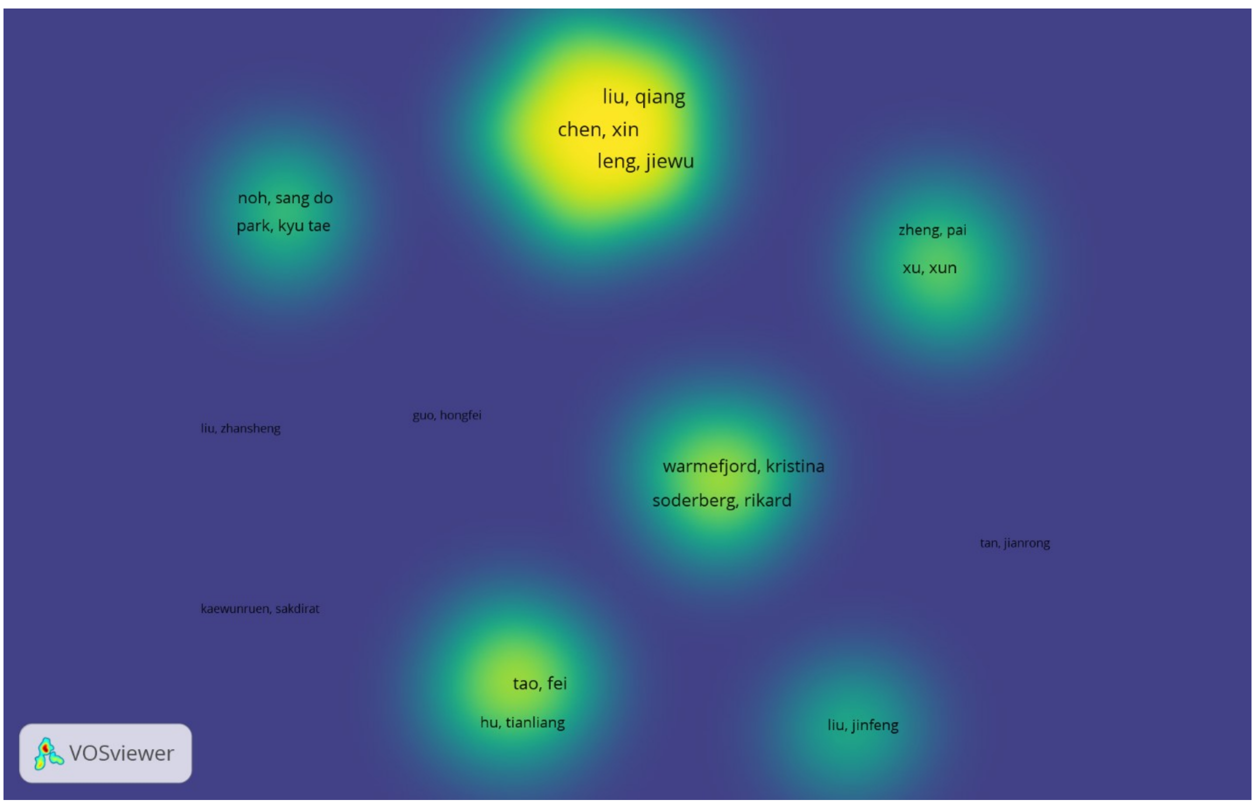

Figure 9: Author co-authorship density visualisation of the literature on digital twin technology, architecture, and applications (n = 3410 authors in the co-authorship network; threshold of 5 documents per author; display of 16 authors).

4.1.5. Co-occurrence of Author Keywords

Keywords are important words or phrases of an article, that usually represent its main content (Chen, et al., 2016). Co-occurrence of these keywords exhibit their interconnectedness based on their combined presence in articles. The co-occurrence of author keywords for this study are shown in Figure 10 with a network visualisation. For inclusion purposes, the minimum number of occurrences of keywords was set to 10. Of the 3110 keywords identified, 25 met the threshold. The total number of occurrences and total link strength for the keywords were calculated and the most frequent keyword was Digital Twin, with 542 occurrences and a total link strength of 429. The next keyword was Industry 4.0 with 89 occurrences and total link strength of 137, and then Internet of Things with 60 occurrences and a total link strength of 111. The occurrences and total link strength of the remaining keywords are presented in the format (X, Y) in that order. The next keyword was Cyber-Physical Systems with (49, 77), then Smart Manufacturing with (40, 63), Simulation with (46, 58), Machine Learning with (43, 58), Artificial Intelligence with (27, 45), Manufacturing (18, 37), and then Virtual Reality with (22, 32). These are the top 10 keywords based on the occurrences and total link strength.

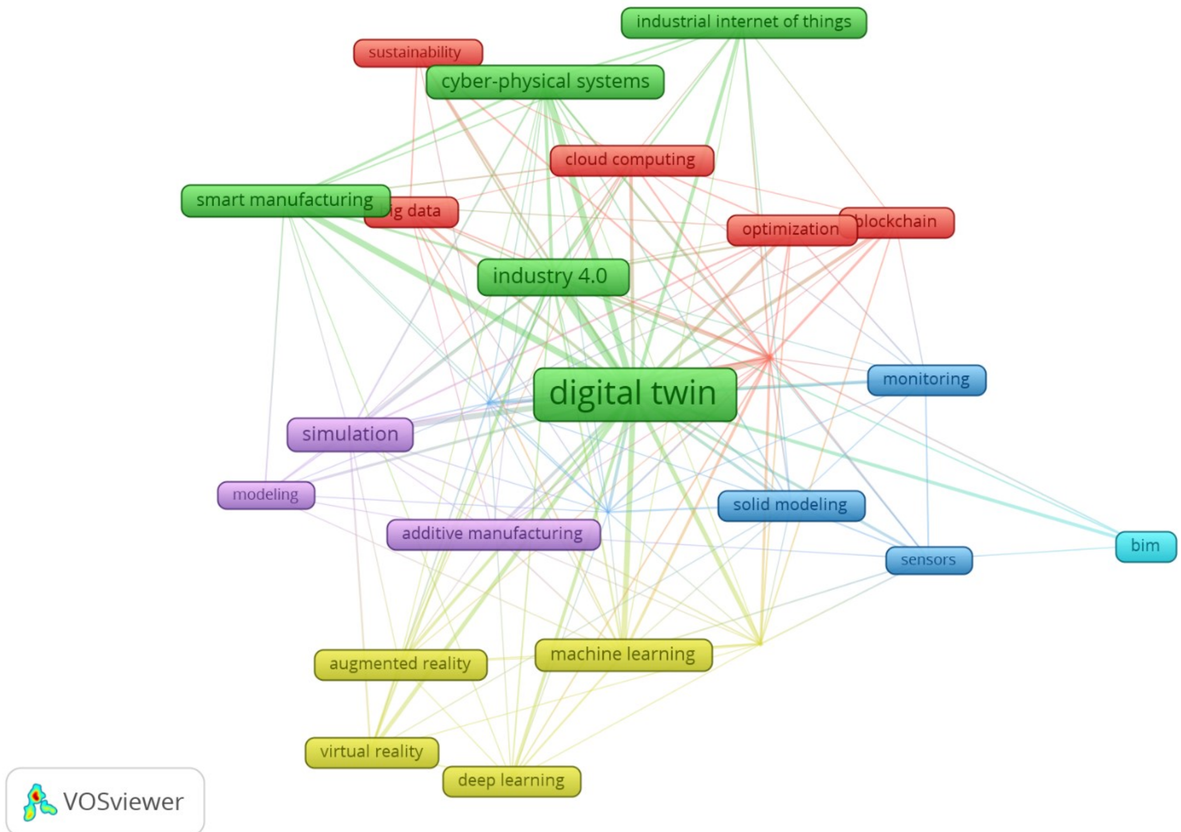

Figure 10: Co-occurrence of author keywords network visualisation map for literature in digital twin technology, architecture, and application until May 31, 2021 (n=3110 keywords; threshold of 10 occurrences; display of 21 keywords).

Considering Figure 10, the colours of the nodes represented as frames signify different clusters in which the keywords are regularly linked to each other. The cluster with the green colour consists of Digital twins, Industry 4.0, Cyber-Physical Systems, Industrial Internet of Things, and Smart Manufacturing, four of which are among the top 5 keywords.

4.1.6. Citation and Co-citation Analyses

Citation and co-citation analysis were employed in order to identify the most influential authors publishing on digital twin technology, its architecture, and applications. The citation analysis, using the Web of Science citation of authors, highlights the most prominent authors such as Fei Tao with 1266 citations, A.Y.C. Nee with 653 citations, Qinglin Qi with 617 citations, and the others, making up the top 20 as shown in Table 2.

Table 2: Top 20 most cited authors in research on digital twin technology, architecture, and applications until May 31, 2021.

Rank	Author	Number of Documents	Web of Science Citations	Citations Per Document
1	Fei Tao	11	1266	115
2	A.Y.C. Nee	6	653	109
3	Qinglin Qi	5	617	123
4	Meng Zhang	4	524	131
5	Qiang Liu	8	475	59
6	Xin Chen	7	461	66
7	Jiewu Leng	7	461	66
8	Ang Liu	4	414	104
9	Hao Zhang	5	394	79
10	Ding Zhang	5	331	66
11	Nabil Anwer	4	325	81
12	Xun Xu	8	308	39
13	Benjamin Schleich	4	282	71
14	Sandro Wartzack	2	276	138
15	Luc Mathieu	1	276	276
16	Morteza Ghobakhloo	1	244	244
17	Douxi Yan	5	240	48
18	He Zhan	1	222	222
19	T. DebRoy	3	220	73
20	Abdulmotaleb El Saddik	4	210	53

The outcome of the author co-citation analysis is presented in Table 3. From the table, the top 20 most influential authors identified in terms of co-citations include 5 of the most influential authors identified using direct citations: Fei Tao, Qinglin Qi, Benjamin Schleich, Jiewu Leng, and Hao Zhang. Despite this not being an unusual occurrence (Philip & Chatpinyakoop, 2019), it should be noted that Web of Science data includes only the first author of a cited document; other authors are not considered in a co-citation analysis of cited authors, and this could be the explanation for the considerable difference.

Table 3: Top 20 most co-cited authors in research on digital twin technology, architecture, and applications until May 31, 2021.

Rank	Author	Co-citations	Total Link Strength
1	*Fei Tao	821	4768
2	Michael Grieves	281	1841
3	*Qinglin Qi	161	1280
4	Jay Lee	161	1033
5	*Benjamin Schleich	138	969
6	*Jiewu Leng	137	1050
7	Elisa Negri	123	912
8	Thomas H.-J. Uhlemann	117	939
9	Stefan Boschert	117	863
10	Roland Rosen	110	880
11	Edward Glaessgen	106	803
12	Yuqian Lu	102	742
13	Rikard Soderberg	99	725
14	Kazi Masudul Alam	80	635
15	Eric J. Tuegel	79	659
16	Werner Kritzinger	78	571
17	*Hao Zhang	76	707
18	Yingfeng Zhang	67	578
19	Michael Schluse	65	572
20	Cunbo Zhuang	64	585

*Indicates that the author also appeared in **Table 2**.

A citation analysis was performed to complement the identification of the job journals in digital twin research. Based on the number of citations, the top 20 journals were selected. Out of the top 20 most cited journals in Table 4, 17 of the journals are among the top 20 journals ranked according to number of publications in Figure 4. IEEE Access came out as the most prominent journal in both cases.

Table 4: Top 20 most cited journals in the publication of research on digital twin technology, architecture, and applications until May 31, 2021.

Rank	Journal	Total Number of Publications	Total Number of citations
1	*IEEE Access	63	1489
2	*CIRP Annals - Manufacturing Technology	18	699
3	*International Journal of Production Research	18	566
4	*Journal of Manufacturing Systems	44	539
5	*Robotics and Computer-Integrated Manufacturing	17	353
6	*IEEE Transactions on Industrial Informatics	7	351
7	*Journal of Ambient Intelligence and Humanized Computing	7	280
8	*The International Journal of Advanced Manufacturing	23	245
9	*International Journal of Computer Integrated Manufacturing	22	245
10	*Journal of Cleaner Production	11	219
11	*Sustainability-Basel	24	169
12	*Applied Sciences-BASEL	46	166
13	Computers in Industry	5	158
14	*Journal of Management in Engineering	10	118
15	*Journal of Intelligent Manufacturing	11	110
16	*Automation in Construction	11	107
17	*SENSORS-BASEL	29	87
18	*Engineering Fracture Mechanics	9	52
19	IEEE Transactions on Power Electronics	5	46
20	*Energies	14	39

*Indicates that journal appeared in Figure 4.

Similarly, a citation analysis was conducted to identify the most influential articles in digital twin literature. For inclusion, the minimum number of citations per document was set to 80 citations. Out of the 938 documents, 20 met the threshold. Majority of the top 20 most cited articles in Table 5 were authored by the most influential authors presented in Table

2 and Table 3 such as Fei Tao, Qinglin Qi, Benjamin Schleich, and Michael Schluse among others. The Web of Science citation counts of the topmost cited articles on digital twins are in moderation, as compared to citation counts of other digital technologies.

Table 5: Top 20 most cited articles on digital twin technology, architecture, and applications until May 31, 2021.

Rank	**Article**	**Article Title**	**Times Cited, WoS Core**
1	Qi (2018)	Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison	288
2	Schleich (2017)	Shaping the digital twin for design and production engineering	276
3	Ghobakhloo (2018)	The future of manufacturing industry: a strategic roadmap toward Industry 4.0	244
4	Tao (2017)	Digital Twin Shop-Floor: A New Shop-Floor Paradigm Towards Smart Manufacturing	244
5	Tao (2019c)	Digital Twin in Industry: State-of-the-Art	222
6	Alam (2017)	C2PS: A Digital Twin Architecture Reference Model for the Cloud-Based Cyber-Physical Systems	197
7	Soderberg (2017)	Toward a Digital Twin for real-time geometry assurance in individualized production	150
8	Tao (2019b)	Digital twin-driven product design framework	138
9	Tao (2018)	Digital twin driven prognostics and health management for complex equipment	128
10	Zhuang (2018)	Digital twin-based smart production management and control framework for the complex product assembly shop-floor	127
11	Bolton (2018)	Customer experience challenges: bringing together digital, physical, and social realms	116
12	Zheng (2018)	A systematic design approach for service innovation of smart product-service systems	115

13	Zhang (2017)	A Digital Twin-Based Approach for Designing and Multi-Objective Optimization of Hollow Glass Production Line	112
14	Liu (2019)	Digital twin-driven rapid individualised designing of automated flow-shop manufacturing system	109
15	Ivanov (2021)	A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0	107
16	Tao (2019a)	Digital Twins and Cyber-Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison	103
17	Leng (2019)	Digital twin-driven manufacturing cyber-physical system for parallel controlling of smart workshop	103
18	Knapp (2017)	Building blocks for a digital twin of additive manufacturing	103
19	Ding (2019)	Defining a Digital Twin-based Cyber-Physical Production System for autonomous manufacturing in smart shop floors	101
20	Schluse (2018)	Experimentable Digital Twins-Streamlining Simulation-Based Systems Engineering for Industry 4.0	87

The articles were grouped into topical clusters and differentiated with colours as shown in Figure 11. Cluster 1 is made up of 6 red coloured frames. It is made up of Alam (2017), Soderberg (2017), Zhang (2017), Liu (2019), Leng (2019), and Ding (2019). These articles concentrated on digital twin architecture and digital twin design methodologies. Cluster 2, consisting of 5 green coloured frames focused on digital twins in manufacturing to achieve smart manufacturing. This cluster consisted of Qi (2018), Schleich (2017), Ghobakhloo (2018), Tao (2018), and Tao (2019a), three of which are the top 3 most cited articles. Cluster 3 consists of 4 blue frames: Tao (2019c), Tao (2019b), Knapp (2017), and Schluse (2018). These articles focused on emerging and fast-growing technologies in digital twins. Cluster 4, making up of Tao (2017) and Zhuang (2018) in yellow frames focused on digital twin shop floors.

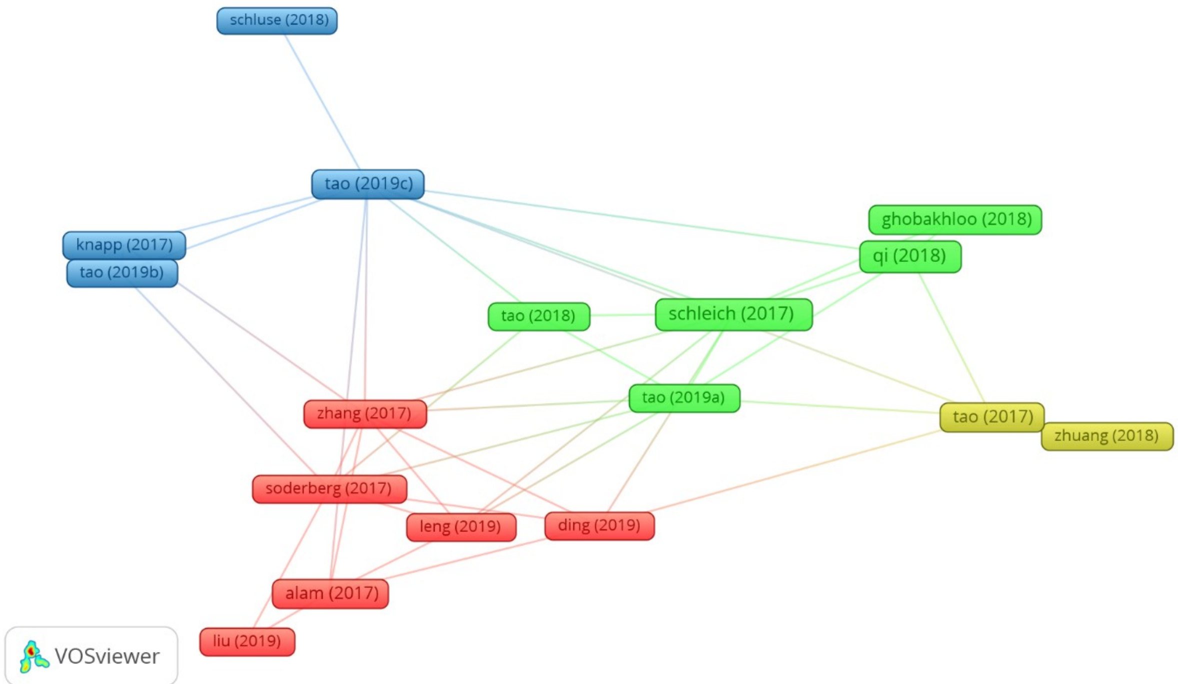

Figure 11: Network visualisation map of the top 20 most cited articles on digital twin technology, architecture, and applications until May 31, 2021.

As the document citation analysis conducted was from only one database, Web of Science, a document co-citation analysis was conducted to obtain a broader perspective of the documents that have been influential in the development of digital twin literature. It was noteworthy that the topmost co-cited document in Table 6 did not appear in the most cited publications in Table 5. Further perusal found that the article was filtered out as a Proceedings Paper during the initial screening of the dataset. There is a reasonable level of overlapping between the top 20 documents with the most citations in Table 5 and the top 20 most co-cited documents in Table 6. There were 9 documents that appeared on both lists: Schleich (2017), Tao (2017), Tao (2019c), Qi (2018), Soderberg (2017), Alam (2017), Tao (2019b), Tao (2018), Zhang (2017). Other documents appearing on the most co-cited documents list is a demonstration of the ability of co-citation analysis to identify influential documents without restriction of a particular database used.

Table 6: Top 20 most co-cited articles on digital twin technology, architecture, and applications until May 31, 2021.

Rank	Article	Article Title	Citations

1	Tao (2018b)	Digital twin-driven product design, manufacturing, and service with big data	176
2	*Schleich (2017)	Shaping the digital twin for design and production engineering	118
3	Grieves (2017)	Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems	107
4	Rosen (2015)	About The Importance of Autonomy and Digital Twins for the Future of Manufacturing	106
5	*Tao (2017)	Digital Twin Shop-Floor: A New Shop-Floor Paradigm Towards Smart Manufacturing	102
6	*Tao (2019c)	Digital Twin in Industry: State-of-the-Art	100
7	*Qi (2018)	Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison	99
8	Glaessgen (2012)	The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles	88
9	Negri (2017)	A Review of the Roles of Digital Twin in CPS-based Production Systems	88
10	*Soderberg (2017)	Toward a Digital Twin for real-time geometry assurance in individualized production	79
11	Kritzinger (2018)	Digital Twin in manufacturing: A categorical literature review and classification	77
12	*Alam (2017)	C2PS: A Digital Twin Architecture Reference Model for the Cloud-Based Cyber-Physical Systems	74
13	Boschert (2016)	Digital Twin—The Simulation Aspect	74
14	Uhlemann (2017)	The Digital Twin: Realizing the Cyber-Physical Production System for Industry 4.0	74
15	*Tao (2019b)	Digital twin-driven product design framework	69
16	*Tao (2018)	Digital twin driven prognostics and health management for complex equipment	66

17	Lee (2015)	A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems	65
18	Grieves (2014)	Digital twin: manufacturing excellence through virtual factory replication	63
19	Lu (2020)	Digital Twin-driven smart manufacturing: Connotation, reference model, applications, and research issues	62
20	*Zhang (2017)	A Digital Twin-Based Approach for Designing and Multi-Objective Optimization of Hollow Glass Production Line	62

*Indicate that the article appeared in Table 5.

It should be noted that the articles are represented by the surname of the first authors in this study only because that it how it was presented by the VOSviewer visualisation.

4.2. Discussion

In this study, the publications on digital twin technology, digital twin architecture and digital twin architecture were retrieved from Web of Science database. This dataset was filtered, analysed, and visualised using descriptive methods and bibliometric methods. Excel, Tableau and the VOSviewer software were used to evaluate and visualise the data. In this circumstance, the publication trend, publication outlets, research areas, co-authorship of countries, organisations, and authors, co-occurrence of author keywords, and citation and co-citation of authors were analysed and presented.

For the publication trend, there is an upward trend in the number of publications on digital twin technology, its architecture and application over the years. The number of publications gained momentum at about 2017 when researchers and industries began to get more curious about the possibilities of digital twins. According to Datta (2017), digital twins is gaining thrust because its possibilities are endless, and it may offer real-time precision. Considering that January to May 31, 2021, have as many publications as the whole of 2020 is an indication that the research and applications of digital twins is growing and will continue to grow speedily.

Considering publication outlets, the journal analysis performed concluded that the research on digital twins, its architecture and applications are being published in good quality multi-disciplinary journals. These journals specialise in engineering, manufacturing, medicine and health, science and technology, robotics, computing, environment, culture, economics, and social sustainability among others. The first journal was IEEE Access. Applied Sciences came in second, then Journal of Manufacturing Systems, then Sensors, and Sustainability, making up the top 5 journals. Using the direct citation analysis in Table 4 it was found that the top 20 most influential journals publishing articles on digital twins, its architecture, and applications in terms of citations and the top 20 in terms of number of publications Figure 4 are in correspondence and are mutually reinforcing. The 3 journals which were not part of the top 20 most influential journals in terms of citations were part of the top 25. Also, the publishers of the journals identified in the studies were ranked in terms of the number of documents published. It was found that the major publishers such as MDPI, IEEEE, Elsevier, Taylor & Francis Ltd, Springer, Wiley, etc are highly ranked publishers. From Figure 5 showing the map of publisher countries, majority of the publications were done in the United Kingdom, USA, Switzerland, and Netherlands. This however does not imply that these countries are the leading countries in relation to research on digital twins.

Several research areas in the form of clusters were identified in the analysis. As no specific research areas were given, the ones provided by Web of Science is what was used. Engineering was the most researched field as it had the most documents. Hartmann & Auweraer (2021) explained that due to the incremental nature of the complexity of engineering design methods, research and development efforts are being made regularly to find easier and more efficient ways, and the research and employment of digital twins is a forward leap. Several of the other research areas identified had Engineering as part of the cluster. The other top research areas sciences such as Computer Science, Management Science, Chemistry, Material Science, among others. As concluded by Ante (2021), digital twin is being considered in several scientific subjects, and this goes to support the claim. Despite Business and Economics making it to the list of top research areas, it was only social science among the top 50 research areas, and it had few publications as compared to the other sciences.

Co-authorship analysis of countries, organisations and authors was performed. Country co-authorships indicated that China had the most co-authorships with a total link strength of 100, 228 documents out of the 938, and 3495 citations. Over the years, China has become a drive to acknowledge in digital technologies. According to Wang, et al. (2017), it is among the first 3 countries for venture capital investment in digital technologies such as 3D printing, artificial intelligence, and virtual reality. It is therefore not startling that majority of the research on digital twins, its architecture and applications is from China, and the country is the strongest at making collaborations with other countries in this field. United States of America (USA), England, Germany, and France, all which are technologically advanced countries, followed in that order. Considering the rankings by Wood (2021), all top 5 countries for co-authorships in digital twin research are among the top 20 most innovative and research-enthusiastic countries. Aside identifying the top organisations and authors in co-authorships of digital twin research, the organisation and author co-authorship analysis confirmed the results of the country co-authorship analysis. Nine out of the eleven organisations shown in Figure 8 are in China. The topmost organisation, University of Hong Kong, is under the Special Administrative Region of China. Beihang University, Jinan University, and National University of Singapore are all institutions in China. The top 5 authors in co-authorship are also from China, even though they may not be currently living in China: Qiang Liu, Xin Chen, Jiewu Leng, Ding Zhang, and Douxi Yan. The findings suggest that research on digital twins are poorly dispersed among countries worldwide. As these countries, organisations and authors are the most influential in co-authorships for this research, it will be prudent for upcoming or not so prominent countries, organisations, and others in terms of research on digital twins to seek for research collaborations with them for a wider audience and a higher impact.

The outcome of the co-occurrence of author keywords signify that researchers have mainly studied these concepts related to the main concept of digital twins: industry 4.0, cyber-physical systems, smart manufacturing, internet of things, industrial internet of things, cloud computing, big data, simulation, augmented reality, virtual reality, machine learning, deep learning, additive manufacturing, sensors, monitoring, optimisation, sustainability, BIM (Building information modelling), and blockchain. Digital twins, being at the core of industry 4.0, incorporates digital technologies like internet of things, machine learning, cloud computing, big data, etc to create cyber-physical systems, virtual realities, and augmented realities via simulations and modelling. Digital twins facilitate with real-time monitoring for optimisation and sustainability. Also, application of digital twins in manufacturing, known as smart manufacturing, is noticeably high as put forth by (Fuller, et al., 2020).

An author citation analysis was undertaken to rank the most influential authors in the research of digital twins. The top 20 most influential authors are presented in Table 2 with Fei Tao being ranked number 1. Among the top authors presented in terms of citations, majority, including Fei Tao, Qiang Liu, Xin Chin, Xu Xun, Douxi Yan, and A.Y.C. Nee were part of the top authors in co-authorships in Figure 9. An author co-citation analysis was undertaken to complement the author citation analysis, and although the results were not greatly overlapping, Fei Tao remained the most influential author, while Jiewu Leng, Qinglin Qi, Hao Zhang, and Benjamin Schleich made it as part of the top 20 authors for co-citation (Refer to Table 3). The results also highlighted a major gender bias in the research of digital twins. This goes to support the assertion by García-González, et al. (2019) that there is a gender inequality in research. The results of the citation analysis for journals was in conformity with the results of the journal ranking using number of publications. Although there were changes in the rankings, 17 of the journals from the list of top 20 journals using number of publications in Figure 4 appeared in the list of top 20 journals using direct citations in Table 4.

From the document citation and co-citation analysis, the most influential articles in digital twin research were identified in Table 5 and Table 6. The articles are grouped into clusters that focused on smart manufacturing, digital twin architecture and design methods, emerging technologies in digital twins such as product design and additive manufacturing, and digital twin shop floors. These articles were authored by majority of the top authors identified in Table 2 and Table 3 and were also published in the most influential journals.

5. Conclusion and Recommendation

5.1. Conclusion

In the past few years, digital twin research has increased significantly. The concept has impacted several sectors such as manufacturing, production, computing, engineering, etc. According to the analysis, it can be concluded that, digital twin research will continue to increase in the years to come, especially in the engineering, computing, and other applied sciences. However, there is more room for further research in the business field, and other social sciences. Also, the outcome of this research is a wakeup call for countries and organisations trailing in digital twin research, and researchers, particularly females to invest more in digital twin research, since studies have shown that digital twin is a worthwhile digital technology with great prospects for effectiveness, efficiency, and optimisation.

5.2. Limitations

The limitations of this study include the use of only one index for the identification of publications for the bibliometric analysis. Although Web of Science is a major index for bibliometric analysis, there is the possibility of exclusion of key documents in the research of digital twins as these studies may not be available in the index. Also, since this is a quantitative study, it is not possible to identify the full impact of publications as a qualitative study would have. For instance, an author may not be part of the top authors in terms of number of documents but may be very influential in his field of study.

5.3. Future Works

It is recommended that in future, several indices be combined to replicate this study for a broader spectrum of publications. It is also recommended that a qualitative study of the research in digital twins be performed, as a qualitative study of the research publications in digital twins may provide further information on the outcome of this study. A bibliometric analysis for digital twin research can also be performed in the different research areas in order to identify the key authors, publishers, documents, and journals in these fields of study.

References

Abburu, S. et al., 2020. *COGNITWIN – Hybrid and Cognitive Digital Twins for the Process Industry.* Cardiff, UK, s.n., pp. 1-8.

Agostino, Í. R. S., Broda, E., Frazzon, E. M. & Freitag, M., 2020. Using a digital twin for production planning and control in industry 4.0. In: *Scheduling in Industry 4.0 and Cloud Manufacturing.* Cham: Springer, pp. 39-60.

Ahmadi, A., Cherifi, C., Cheutet, V. & Ouzrout, Y., 2017. *A Review of CPS 5 Components Architecture for Manufacturing Based on Standards.* Colombo, Sri Lanka, IEEE Xplore.

Aivaliotis, P., Georgoulias, K. & Ghryssolouris, C., 2019. The use of Digital Twin for predictive maintenance in manufacturing. *International Journal of Computer Integrated Manufacturing,* 32(11), pp. 1067-1080.

Al-Ali, A. R. et al., 2020. Digital Twin Conceptual Model within the Context of Internet of Things. *Future Internet,* 12(163).

Alam, K. M. & Saddik, A. E., 2017. C2PS: A Digital Twin Architecture Reference Model for the Cloud-Based Cyber-Physical Systems. *IEEE Access,* Volume 5, pp. 2050-2062.

Amisha, Malik, P., Pathania, M. & Rathaur, V. K., 2019. Overview of artificial intelligence in medicine. *Journal of Family Medicine and Primary Care,* 8(7), p. 2328–2331.

Ante, L., 2021. Digital twin technology for smart manufacturing and industry 4.0: A bibliometric analysis of the intellectual structure of the research discourse. *Manufacturing Letters,* pp. 96-102.

Antony, J., Psomas, E., Garza-Reyes, J. A. & Hines, P., 2020. Practical implications and future research agenda of lean manufacturing: a systematic literature review. *Production Planning & Control.*

Ardito, L., Scuotto, V., Giudice, M. D. & Petruzzelli, A. M., 2019. A bibliometric analysis of research on Big Data analytics for business and management. *Management Decision,* 57(8), pp. 1993-2009.

Arunachalam, D., Kumar, N. & Kawalek, J. P., 2018. Understanding big data analytics capabilities in supply chain management: Unravelling the issues, challenges and implications for practice. *Transportation Research Part E: Logistics and Transportation Review,* Volume 114, pp. 416-436.

Barricelli, B. R., Casiraghi, E. & Fogli, D., 2019. A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. *IEEE Access,* Volume 7, pp. 167653-167671.

Barricelli, B. R., Casiraghi, E. & Fogli, D., 2019. A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. *IEEE Access,* Volume 7, pp. 167653 - 161671.

Bauer, M. et al., 2013. IoT Reference Model. In: *Enabling Things to Talk.* Berlin, Heidelberg: Springer, pp. 113-162.

Bellamy III, W., 2018. *Aviation Today.* [Online]
Available at: https://www.aviationtoday.com/2018/09/14/boeing-ceo-talks-digital-twin-era-aviation/
[Accessed 26 July 2021].

Ben-Daya, M., Hassini, E. & Bahroun, Z., 2019. Internet of things and supply chain management: a literature review. *International Journal of Production Research,* 57(15-16), pp. 4719-4742.

Bramer, W. M., Rethlefsen, M. L., Kleijnen, J. & Franco, O. H., 2017. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Systematic Reviews,* 6(245).

Celebi, D., 2020. Bibliometric analysis of social entrepreneurship in gastronomy tourism. *Tourism: An International Interdisciplinary Journal,* 68(1), pp. 58-67.

Centomo, S., Dall'Ora, N. & Fummi, F., 2020. *The Design of a Digital-Twin for Predictive Maintenance.* Vienna, Austria, IEEE, pp. 1781-1788.

Chahrour, M. et al., 2020. A Bibliometric Analysis of COVID-19 Research Activity: A Call for Increased Output. *Cureus,* 12(3).

Chen, C., 2019. *How to use CiteSpace.* Victoria: Leanpub.

Chen, H.-M., Kazman, R., Haziyev, S. & Hrytsay, O., 2015. *Big Data System Development: An Embedded Case Study with a Global Outsourcing Firm.* Florence, Italy, IEEE.

Chen, X. et al., 2016. Mapping the research trends by co-word analysis based on keywords from funded project. *Procedia Computer Science,* Volume 91, pp. 547-555.

Ciano, M. P., Pozzi, R., Rossi, T. & Strozzi, F., 2020. Digital twin-enabled smart industrial systems: a bibliometric review. *International Journal of Computer Integrated Manufacturing,* pp. 1-19.

Clark, J., 2016. *IBM Business Operations Blog.* [Online]
Available at: https://www.ibm.com/blogs/internet-of-things/what-is-the-iot/
[Accessed 13 June 2021].

Crawford, M., 2021. *The American Society of Mechanical Engineers (ASME).* [Online]
Available at: https://www.asme.org/topics-resources/content/7-digital-twin-applications-for-manufacturing
[Accessed 11 September 2021].

Cui, Y., Kara, S. & Chan, K. C., 2020. Manufacturing big data ecosystem: A systematic literature review. *Robotics and Computer-Integrated Manufacturing,* Volume 62, p. 101861.

Daneshkhah, A., Hosseinian-Far, A. & Chatrabgoun, O., 2017. Sustainable Maintenance Strategy Under Uncertainty in the Lifetime Distribution of Deteriorating Assets. In: *Strategic Engineering for Cloud Computing and Big Data Analytics.* Cham: Springer, pp. 29-50.

Datta, S. P. A., 2017. Emergence of Digital Twins. *Journal of Innovation Management,* Volume 5, pp. 14-34.

Dave, A., 2020. *iotforall.com.* [Online]
Available at: https://www.iotforall.com/how-digital-twins-accelerate-the-growth-of-iot
[Accessed 24 July 2021].

Dilmegani, C., 2021. *AI Multiple.* [Online]
Available at: https://research.aimultiple.com/digital-twins/
[Accessed 24 July 2021].

Dohrmann, K., Gesing, B. & Ward, J., 2019. *Digital Twins in Logistics: A DHL perspective on the impact of digital twins on the logistics industry.* Troisdorf: DHL Customer Solutions & Innovation.

Ellegaard, O. & Wallin, J. A., 2015. The bibliometric analysis of scholarly production: How great is the impact?. *Scientometrics,* 105(3), pp. 1809 - 1831.

Elsevier, 2021. *Elsevier.* [Online]
Available at: https://www.elsevier.com/en-gb/solutions/scopus
[Accessed 09 August 2021].

Esmaeilbeigi, M. et al., 2020. A low cost and highly accurate technique for big data spatial-temporal interpolation. *Applied Numerical Mathematics,* Volume 153, p. 492–502.

Farsi, M., Daneshkhah, A., Hosseinian-Far, A. & Jahankhani, H., 2020. *Digital Twin Technologies and Smart Cities.* Cham: Springer.

Frankenfield, J., 2020. *Investopedia.* [Online]
Available at: https://www.investopedia.com/terms/d/data-analytics.asp
[Accessed 30 December 2020].

Fuller, A., Fan, Z., Day, C. & Barlow, C., 2020. Digital Twin: Enabling Technologies, Challenges and Open Research. *IEEE Access,* Volume 8, pp. 108952-108971.

García-González, J., Forcén, P. & Jimenez-Sanchez, M., 2019. Men and women differ in their perception of gender bias in research institutions. *PLoS ONE,* 14(12).

Gelernter, D. H., 1991. *Mirror worlds, or, The day software puts the universe in a shoebox--: how it will happen and what it will mean.* New York: Oxford University Press.

Google Scholar, 2021. *Google Scholar.* [Online]
Available at: https://scholar.google.co.uk/intl/en/scholar/about.html
[Accessed 24 May 2021].

Greer, C., Burns, M., Wollman, D. & Griffor, E., 2019. *Cyber-Physical Systems and Internet of Things.* s.l.:National Institute of Standards and Technology (NIST) Special Publication 1900-202.

Grieves, M., 2014. *Digital Twin: Manufacturing Excellence Through Virtual Factory Replication.* s.l.:Michael W. Grieves, LLC.

Grieves, M. & Vickers, J., 2017. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In: *Kahlen FJ., Flumerfelt S., Alves A. (eds) Transdisciplinary Perspectives on Complex Systems.* Cham: Spinger, pp. 85-113.

Grieves, M. & Vickers, J., 2017. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In: F. Kahlen, S. Flumerfelt & A. Alves, eds. *Transdisciplinary Perspectives on Complex Systems.* ham: Springer, pp. 85-113.

Hallinger, P. & Chatpinyakoop, C., 2019. A Bibliometric Review of Research on Higher Education for Sustainable Development, 1998–2018. *Sustainability,* 11(2401).

Hartmann, D. & Auweraer, H. V. d., 2021. Digital Twins. In: *Progress in Industrial Mathematics: Success Stories.* Cham: Springer, pp. 3-17.

Hinduja, H., Kekkar, S., Chourasia, S. & Chakrapani, H. B., 2020. Industry 4.0: Digital Twin and its Industrial Applications. *International Journal of Science, Engineering and Technology,* 8(4).

Hosseinian-Far, A., Ramachandran, M. & Slack, C. L., 2018. Emerging Trends in Cloud Computing, Big Data, Fog Computing, IoT and Smart Living. In: *Technology for Smart Futures.* Cham: Springer, pp. 29-40.

Huang, L., Chen, K. & Zhou, M., 2020. Climate change and carbon sink: a bibliometric analysis. *Environmental Science and Pollution Research,* Volume 27, p. 8740–8758.

Huynh, B. H., Akhtar, H. & Sett, M. K., 2019. *A Universal Methodology to Create Digital Twins for Serial and Parallel Manipulators.* Barl, Italy, 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), pp. 3104-3109.

IEEE Big Data, 2018. *Big Data For Digital Twins.* Seattle, IEEE.

IFS, 2021. *Digital twins in Commercial Aviation.* s.l.:IFS.

Jalal, S. K., 2019. Co-authorship and co-occurrences analysis using BibliometrixR package: a case study of India and Bangladesh. *Annals of Library and Information Studies,* Volume 66, pp. 57-64.

Jones, D. et al., 2020. Characterising the Digital Twin: A systematic literature review. *CIRP Journal of Manufacturing Science and Technology,* 29(Part A), pp. 36-52.

Josifovska, K., Yigitbas, E. & Engels, G., 2019. *Reference Framework for Digital Twins within Cyber-Physical Systems.* Montreal, QC, Canada, s.n., pp. 25-31.

Kantarci, A., 2021. *AI Multiple.* [Online]
Available at: https://research.aimultiple.com/digital-twin-applications/
[Accessed 19 June 2021].

Karakus, M., Ersozlu, A. & Clark, A. C., 2019. Augmented Reality Research in Education: A Bibliometric Study. *EURASIA Journal of Mathematics, Science and Technology Education,* 15(10).

Kasper, L., Bacher, T., Birkelbach, F. & Hofmann, R., 2020. Digitalization possibilities and the potential of the Digital Twin for steam supply systems. *VGB PowerTech,* Volume 11, pp. 45-52.

Ketzler, B. et al., 2020. Digital Twins for Cities: A State of the Art Review. *Built Environment,* 46(4), pp. 547-573.

Ketzler, B. et al., 2020. Digital Twins for Cities: A State of the Art Review. *Built Environment,* 46(4), pp. 547-573(27).

Khajavi, S. H. et al., 2019. Digital Twin: Vision, Benefits, Boundaries, and Creation for Buildings. *IEEE Access,* Volume 7, pp. 147406 - 147419.

Khan, A. et al., 2020. Towards Smart Manufacturing Using Spiral Digital Twin Framework and Twinchain. *IEEE Transactions on Industrial Informatics.*

Korom, P., 2019. A bibliometric visualization of the economics and sociology of wealth inequality: a world apart?. *Scientometrics,* Volume 118, pp. 849-868.

Kosowatz, J., 2020. *The American Society of Mechanical Engineers.* [Online] Available at: https://www.asme.org/topics-resources/content/top-10-growing-smart-cities [Accessed 26 July 2021].

Kosowatz, J., 2021. *The American Society of Mechanical Engineers (ASME).* [Online] Available at: https://www.asme.org/topics-resources/content/smart-cities-look-for-digital-twins
[Accessed 25 July 2021].

Kritzinger, W. et al., 2018. Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine,* 51(11), pp. 1016-1022.

Kumar, H., Singh, M. K., Gupta, M. P. & Madaan, J., 2020. Moving towards smart cities: Solutions that lead to the Smart City Transformation Framework. *Technological Forecasting and Social Change,* Volume 153.

Kumar, P. R., Raj, H. P. & Jelciana, P., 2018. Exploring Data Security Issues and Solutions in Cloud Computing. *Procedia Computer Science,* Volume 125, p. 691–697.

Lee, I. & Lee, K., 2015. The Internet of Things (IoT): Applications, investments, and challenges for enterprises. *Business Horizons,* 58(4), pp. 431-440.

Lee, J., Bagheri, B. & Kao, H.-A., 2015. A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters,* Volume 3, pp. 18-23.

Leng, J. et al., 2019. Digital twin-driven manufacturing cyber-physical system for parallel controlling of smart workshop. *Journal of Ambient Intelligence and Humanized Computing,* Volume 10, pp. 1155-1166.

Lia, J., Zhou, G. & Zhang, C., 2021. A twin data and knowledge-driven intelligent process planning framework of aviation parts. *International Journal of Production Research,* pp. 1-18.

Liu, Y. et al., 2019. A Novel Cloud-Based Framework for the Elderly Healthcare Services Using Digital Twin. *IEEE Access,* Volume 7, pp. 49088 - 49101.

Lohtander, M. et al., 2018. Micro Manufacturing Unit and the Corresponding 3D-Model for the Digital Twin. *Procedia manufacturing,* Volume 25, pp. 55-61.

Macke, J. & Genari, D., 2019. Systematic literature review on sustainable human resource management. *Journal of Clearner Production,* 208(20), pp. 806-815.

Madni, A. M., 2019. *Exploiting Digital Twin Technology to Teach Engineering Fundamentals and Afford Real-World Learning Opportunities.* Tampa, American Society for Engineering Education.

Madni, A. M., 2019. *Exploiting Digital Twin Technology to Teach Engineering Fundamentals and Afford Real-World Learning Opportunities.* Tampa, American Society for Engineering Education.

Majumdar, P. K., Haider, M. F. & Reifsnider, K., 2013. *Multi-physics response of structural composites and framework for modeling using material geometry.* s.l., 54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, p. 1577.

Maria, G., 2020. *Software Advice.* [Online]
Available at: https://www.softwareadvice.com/resources/what-is-digital-twin-technology/
[Accessed 25 May 2021].

Martín-Martín, A., Thelwall, M., Orduna-Malea , E. & López-Cózar , E. D., 2021. Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage via citations. *Scientometrics,* Volume 126, pp. 871-906.

Mikalef, P., Pappas, I. O., Krogstie, J. & Giannakos, M., 2018. Big data analytics capabilities: a systematic literature review and research agenda. *Information Systems and e-Business Management,* Volume 16, pp. 547-578.

Moher, D. et al., 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine,* 6(7).

Moral-Muñoz, J. A., Herrera-Viedma, E., Santisteban-Espejo, A. & Cobo, M. J., 2020. Software tools for conducting bibliometric analysis in science: An up-to-date review. *El profesional de la información,* 29(1).

Negri, E., Fumagalli, L. & Macchi, M., 2017. A Review of the Roles of Digital Twin in CPS-based Production Systems. *Procedia Manufacturing,* Volume 11, pp. 939-948.

Neto, A. A., Deschamps, F., Silva, E. R. & Lima, E. P., 2020. Digital twins in manufacturing: an assessment of drivers, enablers and barriers to implementation. *Procedia CIRP,* Volume 93, pp. 210-215.

Niaros, V., Kostakis, V. & Drechsler, W., 2017. Making (in) the smart city: The emergence of makerspaces. *Telematics and Informatics,* 34(7), pp. 1143-1152.

Nord, J. H., Koohang, A. & Paliszkiewicz, J., 2019. The Internet of Things: Review and theoretical framework. *Expert Systems with Applications,* Volume 133, pp. 97-108.

Okumus, F. et al., 2019. A Bibliometric Analysis of Lodging-Context Research From 1990 to 2016. *Journal of Hospitality & Tourism Research,* 43(2), pp. 210 - 225.

OPC DataHub, n.d.. *OPC DataHub.* [Online]
Available at: https://opcdatahub.com/WhatIsOPC.html
[Accessed 10 September 2021].

Orcajo, E. M., 2021. *LinkedIn.* [Online]
Available at: https://www.linkedin.com/pulse/6-digital-twin-applications-healthcare-revolution-enrique/
[Accessed 25 July 2021].

Ornes, S., 2016. Core Concept: The Internet of Things and the explosion of interconnectivity. *Proceedings of the National Academy of Sciences,* 113(40), pp. 11059-11060.

Paripooranan, C. S., Vivek, D. C., Abishek, R. & Karthik, S., 2020. *An Implementation of AR Enabled Digital Twins for 3-D Printing.* Chennai, India, 2020 IEEE International Symposium of Smart Electronic Systems (iSES).

Philip , H. & Chatpinyakoop, C., 2019. A Bibliometric Review of Research on Higher Education for Sustainable Development, 1998–2018. *Sustainability,* 11(2401).

Philips, 2018. *philips.co.uk.* [Online]
Available at: https://www.philips.co.uk/healthcare/resources/feature-detail/ultrasound-heartmodel?
[Accessed 25 July 2021].

Pires, F. et al., 2019. *Digital Twin in Industry 4.0: Technologies, Applications and Challenges.* s.l., IEEE, pp. 721-726.

Polat, S., Esen, F. & Bilgic, E., 2019. Analysis of the 5Vs of Big Data in Virtual Travel Organizations. In: *Big Data and Knowledge Sharing in Virtual Organizations.* s.l.:IGI Global, pp. 43-70.

Polini, W. & Corrado, A., 2020. Digital Twin of Composite Assembly Manufacturing Process. *International Journal of Production Research,* 58(17), p. 5238–5252.

Pranckute, R., 2021. Web of Science (WoS) and Scopus: The Titans of Bibliographic Information in Today's Academic World. *Publications,* 9(12).

PRISMA, 2021. *PRISMA.* [Online]
Available at: http://www.prisma-statement.org/
[Accessed 25 May 2021].

Qi, Q. & Tao, F., 2018. Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison. *IEEE Access,* Volume 6, pp. 3585 - 3593.

Qi, Q. et al., 2021. Enabling technologies and tools for digital twin. *Journal of Manufacturing Systems,* 58(Part B), pp. 3-21.

Queirós, A., Faria, D. & Almeida, F., 2017. Strengths and limitations of qualitative and quantitative research methods. *European Journal of Education Studies,* 3(9).

Radanliev, P. et al., 2021. Digital twins: artificial intelligence and the IoT cyber-physical systems in Industry 4.0. *International Journal of Intelligent Robotics and Applications.*

Redelinghuys, A., Basson, A. & Kruger, K., 2019. A Six-Layer Digital Twin Architecture for a Manufacturing Cell. In: T. Borangiu, D. Trentesaux, A. Thomas & S. Cavalieri, eds. *Service orientation in holonic and multi-agent manufacturing. SOHOMA 2018 studies in computational intelligence.* Cham: Springer, pp. 412-423.

Redelinghuys, A. J. H., Basson , A. H. & Kruger, K., 2020. A six-layer architecture for the digital twin: a manufacturing case study implementation. *Journal of Intelligent Manufacturing,* Volume 31, p. 1383–1402.

Rice, D. B. et al., 2016. Are MEDLINE searches sufficient for systematic reviews and meta-analyses of the diagnostic accuracy of depression screening tools? A review of meta-analyses. *Journal of Psychosomatic Research,* Volume 87, pp. 7-13.

Rosen, R., Wichert, G. v., Lo, G. & Bettenhausen, K. D., 2015. *About the Importance of Autonomy and Digital Twins for the Future of Manufacturing.* s.l., IFAC-PapersOnLine, pp. 567-572.

Ross, P. T. & Zaidi, N. L. B., 2019. Limited by our limitations. *Perspectives of Medical Education,* 8(4), p. 261–264.

Rudskoy, A., Ilin, I. & Prokhorov, A., 2021. Digital Twins in the Intelligent Transport Systems. *Transportation Research Procedia,* Volume 54, p. 927–935.

Sambunjak, D., Cumpston, M. & Watts, C., 2017. *Module 1: Introduction to conducting systematic reviews. In: Cochrane Interactive Learning: Conducting an intervention review,* s.l.: Cochrane.

Schleich, B., Anwer, N., Mathieu, L. & Wartzack, S., 2017. Shaping the digital twin for design and production engineering. *CIRP Annals,* 66(1), pp. 141-144.
Schleich, B., Anwer, N., Mathieu, L. & Wartzack, S., 2017. Shaping the digital twin for design and production engineering. *CIRP Annals,* 66(1), pp. 141-144.
Schoonenboom, J. & Johnson, R. B., 2017. How to Construct a Mixed Methods Research Design. *Kolner Z Soz Sozpsychol,* 69(Suppl 2), pp. 107-131.

Schroeder, G. N., Steinmetz, C., Pereira, C. E. & Espindola, D. B., 2016. Digital Twin Data Modeling with AutomationML and a Communication Methodology for Data Exchange. *IFAC-PapersOnLine,* 49(30), pp. 12-16.

Sepasgozar, S. M., 2020. Digital Twin and Web-Based Virtual Gaming Technologies for Online Education: A Case of Construction Management and Engineering. *Applied Sciences,* 10(13).

Serrano, L., Sianes, A. & Ariza-Montes, A., 2019. Using Bibliometric Methods to Shed Light on the Concept of Sustainable Tourism. *Sustainability,* 11(24).

Söderberg, R., Wärmefjord, K., Carlson, J. S. & Lindkvist, L., 2017. Toward a Digital Twin for real-time geometry assurance in individualized production. *CIRP Annals,* 66(1), pp. 137-140.

Steindl, G. et al., 2020. Generic Digital Twin Architecture for Industrial Energy Systems. *Applied Sciences,* 10(24).

Suthar, B., Bongale, A. & Kumar, S., 2021. Three Degrees of Freedom Robotic arm and its Digital Twin using Simulink–A Bibliometric Analysis. *Library Philosophy and Practice (e-journal).*

Tao, F., Qi, Q., Wang, L. & Nee, A. Y. C., 2019. Digital Twins and Cyber–Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison. *Engineering,* 5(4), pp. 653-661.

Tao, F., Zhang, M., Liu, Y. & Nee, A. Y., 2018. Digital twin driven prognostics and health management for complex equipment. *CIRP Annals,* 6(1), pp. 169-172.

Teng, X. & Gong, Y., 2018. Research on Application of Machine Learning in Data Mining. *IOP Conference Series: Materials Science and Engineering,* 392(6).

Toivonen, V., Lanz, M., Nylund, H. & Nieminen, H., 2018. The FMS Training Center - a versatile learning environment for engineering education. *Procedia Manufacturing,* Volume 23, pp. 135-140.

UCL, 2021. *UCL Library Services.* [Online]
Available at: https://library-guides.ucl.ac.uk/web-of-science/citation-analysis
[Accessed 09 August 2021].

University of Exeter, 2021. *University of Exeter LibGuides.* [Online]
Available at: https://libguides.exeter.ac.uk/c.php?g=655923&p=4681459
[Accessed 20 September 2021].

Valk, H. v. d. et al., 2020. *A Taxonomy of Digital Twins.* s.l., Americas Conference on Information Systems.

Vassakis, K., Petrakis, E. & Kopanakis, I., 2018. Big Data Analytics: Applications, Prospects and Challenges. In: G. Skourletopoulos, et al. eds. *Mobile Big Data. Lecture Notes on Data Engineering and Communications Technologies.* Cham: Springer, p. 1.

VOSviewer, n.d.. *VOSviewer.* [Online]
Available at: VOSviewer.com
[Accessed 25 May 2021].

Wang, J. et al., 2019. Digital Twin for Rotating Machinery Fault Diagnosis in Smart Manufacturing. *International Journal of Production Research,* 57(12), p. 3920–3934.

Wang, K. W. et al., 2017. *Digital China: Powering the economy to global competitiveness,* New York: McKingsley and Company.

Wood, T., 2021. *Visual Capitalist.* [Online]
Available at: https://www.visualcapitalist.com/national-innovation-the-most-innovative-countries-by-income/
[Accessed 09 09 2021].

Wright , L. & Davidson , S., 2020. How to tell the difference between a model and a digital twin. *Advanced Modeling and Simulation in Engineering Sciences volume,* 7(13).

Xu, Y., Sun, Y., Liu, X. & Zheng, Y., 2019. A Digital-Twin-Assisted Fault Diagnosis Using Deep Transfer Learning. *IEEE Access,* Volume 7, pp. 19991-19999.

Zerhari, B., Lahcen, A. A. & Mouline, S., 2015. *Big Data Clustering: Algorithms and Challenges.* s.l., Proc. of Int. Conf. on Big Data, Cloud and Applications (BDCA'15).

Zhang, X., Chen, H., Wang, W. & Pablos, P. O. d., 2016. What is the role of IT in innovation? A bibliometric analysis of research development in IT innovation. *Behaviour & Information Technology,* 35(12), p. 1130–1143.

Zhaohui, W. et al., 2021. Review on the Construction and Application of Digital Twins in Transportation Scenes. *Journal of System Simulation,* 33(2), pp. 295-305.

Židek, K. et al., 2020. Digital Twin of Experimental Smart Manufacturing Assembly System for Industry 4.0 Concept. *Sustainability,* 12(9).

Zinghini, F., 2020. *LinkedIn.* [Online]
Available at: https://www.linkedin.com/pulse/reduce-costs-iot-digital-twins-frank-zinghini/
[Accessed 11 September 2021].

Zupic, I. & Čater, T., 2015. Bibliometric Methods in Management and Organisation. *Organisational Research Methods,* Volume 18, p. 429–472.