Digital twin paradigm: A systematic literature review

Concetta Semeraro, Mario Lezoche, Hervé Panetto, Michele Dassisti

Digital Twin Paradigm: A Systematic Literature Review

Concetta Semeraro[*,1,2], Mario Lezoche[2], Hervé Panetto[2],
Michele Dassisti[1]

[1] *Department of Mechanical, Mathematics and Management (DMMM), Politecnico di Bari, Bari- Italy (e-mail: {concetta.semeraro, michele.dassisti} @poliba.it).*

[2] *Université de Lorraine, CNRS, CRAN, Nancy, France (e-mail: {concetta.semeraro, mario.lezoche, herve.panetto}@univ-lorraine.fr)*

Highlights

- A systematic literature review is conducted to explore the main features, research and technical challenges in conceiving and building Digital Twins.
- Topic Modelling Analysis has been implemented to provide an up-to-date picture of the digital twin.
- Formal Concept Analysis (FCA) has been applied to understand the digital twin trends and strategies.

Abstract:

Manufacturing enterprises are facing the need to align themselves to the new information technologies (IT) and respond to the new challenges of variable market demand. One of the key enablers of this IT revolution toward Smart Manufacturing is the digital twin (DT). It embeds a "virtual" image of the reality constantly synchronized with the real operating scenario to provide sound information (knowledge model) to reality interpretation model to draw sound decisions. The paper aims at providing an up-to date picture of the main DT components, their features and interaction problems. The paper aims at clearly tracing the ongoing research and technical challenges in conceiving and building DTs as well, according to different application domains and related technologies. To this purpose, the main questions answered here are: 'What is a Digital Twin?'; 'Where is appropriate to use a Digital Twin?'; 'When has a Digital Twin to be developed?'; 'Why should a Digital Twin be used?'; 'How to design and implement a Digital Twin?'; 'What are the main challenges of implementing a Digital Twin?'. This study tries to answer to the previous questions funding on a wide systematic literature review of scientific research, tools, and technicalities in different application domains.

Keywords: Digital Twin; Industry 4.0; Cyber-Physical Systems; Predictive manufacturing.

Introduction

In the past, due to the lack of information technologies, the physical space played the main role in controlling the production in shop floors, leading to low efficiency, accuracy, and transparency. Until the 20th century, technologies such as computers, simulation tools, Internet, and wireless networks introduced a parallel virtual space to virtualize physical assets and to enable the cooperation with assets remotely. This has provided a possibility to conduct plans and operations more efficiently and

effectively (Tao and Zhang, 2017). Nowadays, with the developments of new generation information technologies (New IT), the integration and the interaction between the physical and virtual spaces is becoming increasingly important. This will create new potentialities for improving the current operating situations and technologies in the fields of design, manufacturing, and service (Büchi et al., 2020). Various countries are converging on this trend as the next industrial revolution (Suh, 1984), (Prasad, 1989), and have proposed related national strategies, such as the "Industry 4.0" in Germany (Kagermann et al., 2013); the "Advanced Manufacturing" or "Smart Manufacturing" in the United States (Yao et al., 2019); the "Society 5.0" in Japan; the "Made in China 2025" in China; the "Industry of the Future" in France; the "Intelligent Factory" in Italy (Osterrieder et al., 2019) and more generally "The Factory of Future" in Europe (Drath and Horch, 2014). Although the strategies are proposed under different environments, their common objective is to capture the opportunity brought by the integration of the physical and virtual spaces (Hermann et al., 2016). The fusion of the physical and virtual spaces is motivated to ensure a better flexibility and scalability of manufacturing systems through information technologies (Dassisti and De Nicolò, 2012), (Pirola et al., 2020). The current digital transformation of enterprises requires the design and application of digital models, called digital twins, which represent a set of knowledge of the real processes (Panetto et al., 2019), (Dassisti et al., 2019a). The digital twin (DT) aims at creating high-fidelity virtual models for each physical entity to emulate their states and behaviours with abilities of evaluating, optimizing, and predicting (Graessler and Poehler, 2017). The concept of using "twins" dated back to NASA's Apollo program, where two identical space vehicles were built to allow mirroring the conditions of the space vehicle during the mission. Professor Grieves at the University of Michigan firstly put forward the concept of 'Digital Twin' in Product Life cycle Management (PLM) courses in 2003 (Grieves and Vickers, 2017). The digital twin refers to a holistic, digital engineering view from the product design and development to production planning, production engineering, production, and associated services (Product Life cycle Management). The DT can be developed for each phase of the product life cycle absolving different functions (Dassisti and Semeraro, 2018). The digital twin in the design stage can help designers to configure and validate more quickly the future scenarios (Brettel et al., 2014). The DT can help decision maker to accurately interpret the market demands and the customer preferences (Semeraro et al., 2019b). At the manufacturing phase, DT may enable the simulation, and thus the decision maker, to analyse the interactive behaviours among production factors by collecting data from order, design, purchase, production planning, manufacturing, and product usage stage. The DT can help optimizing and evaluating in real time the production planning and the behaviour of the production process. At the service stage, DT relies on real time state monitoring and virtual operations such as maintenance to predict the remaining life of components or products (Lee and Kim, 2018). The virtual replication of a physical system is a rather complex task and therefore it requires the availability of a large amount of data and models that represent the modelled system, even though there are not specific criteria to follow (Park et al., 2019).

In the scientific literature reviewed in this paper, several studies have been devoted to analysing the DT concept, which results different as the context of application changes (aerospace, manufacturing, city management). In each context, digital twins have their own specificity within the life cycle phase of the product: namely design, manufacturing, and service. As a result, each application of DT varies depending on a different perspective and needs accordingly. In this context, the paper aims at

providing an advanced and up-to-date picture of the state-of-the-art considering the main features and challenges of existing scientific research on DT's, focussing on the different application domains and their related technologies. The paper funds its scientific basis on information, principles and hints derived from a systematic scientific literature review employing text mining techniques to identify textual patterns, topic modelling, and new insights. Hereof, seven main research questions are raised and discussed: (1) 'What is a Digital Twin?'; (2) 'Where is appropriate to use a Digital Twin?'; (3) 'Who is doing Digital Twins?'; (4) 'When has a Digital Twin to be developed?'; (5) 'Why should a Digital Twin be used?'; (6) 'How to design and implement a Digital Twin?'; (7) 'What are the main challenges of implementing a Digital Twin?'. The research questions concern six main aspects namely: the Digital Twin definition; the application contexts; the life cycle phases; the functions; the architecture and the components, the research challenges. These aspects have been discussed in detail to define and explore the main features, research and technical challenges in conceiving and building Digital Twins. To serve this purpose, Formal Concept analysis (FCA) was run on to get deep into the definition of DT life cycle phases and its functions and into the DT architecture and its components. The outcome is then a multi-perspective picture of the Digital Twin, forming a paradigm emerging from scientific literature.
The content is structured as the following: the literature review methodology is described in section 1, the state of art in section 2, the digital twin paradigm in section 3. The conclusion and the research challenges are argued in section 4.

1 Literature Review Methodology

The literature review focuses on works related to DT technology. A systematic literature search was conducted in the Scopus, Elsevier and ScienceDirect database, covering most of the peer-reviewed interdisciplinary research papers. The methodology applied is composed by three-step approach: Paper selection; Extraction of DT features; Knowledge representation. Each step is described below.

Step 1: Paper Selection: Publications identification and screening

The present study forms a state-of-the-art on digital twin. The focal point of the study was based on DT representation in different scientific papers. This review was conducted based on content analysis. The Scopus, Elsevier and ScienceDirect scientific databases were used to find the literature for this review. In addition to 'Digital Twin', search terms such as 'Factory of Future', 'Industry 4.0 technologies', 'Cyber-physical system', 'Predictive manufacturing' were used to search for suitable papers within the targets and scopes of this review paper. We found over 300 papers from our search. The relevant literature was selected by analysing the title, abstract, keywords, paper contents and journal's main topic of interest. Finally, we selected the papers based on impact factor, citation, and review process. We identified and analysed 150 papers of which 35 in the fields of 'Factory of Future', 'Industry 4.0 technologies', 'Cyber-physical system', 'Predictive manufacturing' and 115 in the field of 'Digital Twin'. The selected papers on the digital twin present the following parameters: Time Span: 2002–2020; Language: English; Type = "Article"; "Journal Paper"; "Conference Proceeding"; "Book Chapter", as recapped in Figure 1.

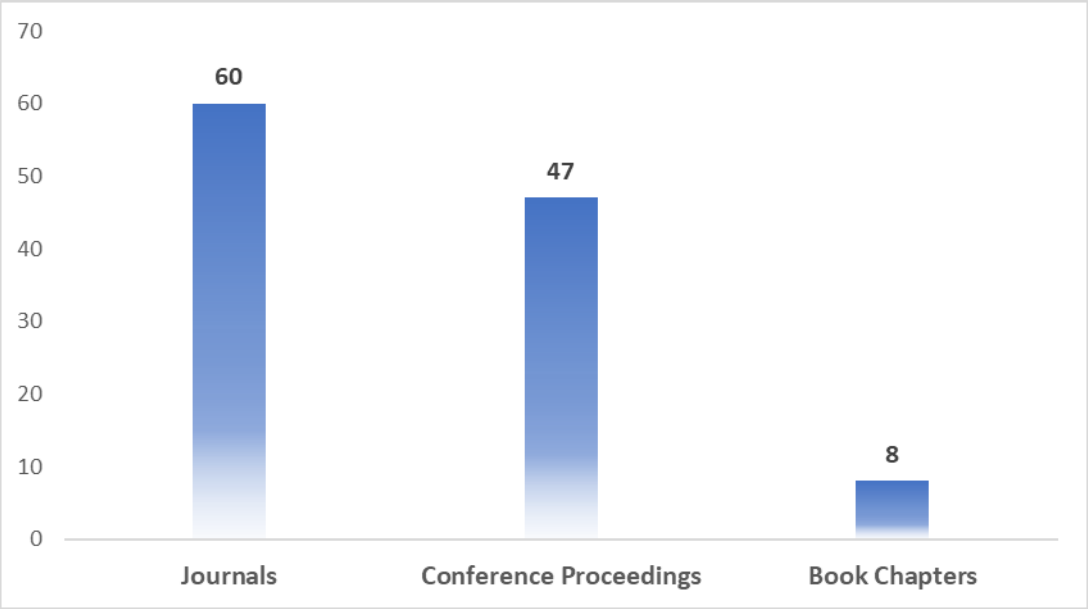


Figure 1: Digital Twin Paper Distribution as type of Bibliographical Reference

Step 2 Extraction of DT Features: Text mining analysis

Following the publications identification and screening, a technical approach has been designed and applied to extract all possible features and information from the selected DT items. Text mining analysis has been selected for this specific purpose. Text mining is the process of analysing text to extract information that can be useful for different purposes (Hearst, 2003). A set of text mining techniques have been used employing Orange tool (Ljubljana, 2005). It allows to design and create workflows by linking predefined or user-designed components called widgets. Two different models were developed in Orange to this aim, as shown in figure 2 and in figure 3, for analysing the selected papers to capture key concepts, trends, and hidden relationships in DT studies.

The first model aims to apply the text mining techniques to DT definitions to clarify what is a digital twin and why it matters. The workflow below shows that all DT definitions were collected, listed, and pre-processed to perform the hierarchical clustering algorithm (HCA) (Grira et al., 2004). HCA is an unsupervised clustering technique that groups similar objects into groups called clusters. The distance between two clusters is computed by the cosine distance because it is a good measure of semantic relatedness (Mikolov et al., 2013). The endpoint is a hierarchy of nested clusters, called dendrogram, where each cluster is distinct from each other cluster, and the objects within each cluster are broadly similar to each other. In this model, the HCA groups the DT definitions in a set of clusters that can be visualized in a data table or in a word cloud respectively named *"Data Table (Clusters DT Definitions)"* and *"Word Cloud of each Cluster"*. The clusters and their respective word clouds are discussed in section 2 to initiate and structure a comprehensive review on the state of art of Digital Twins comparing the definitions provided in literature.

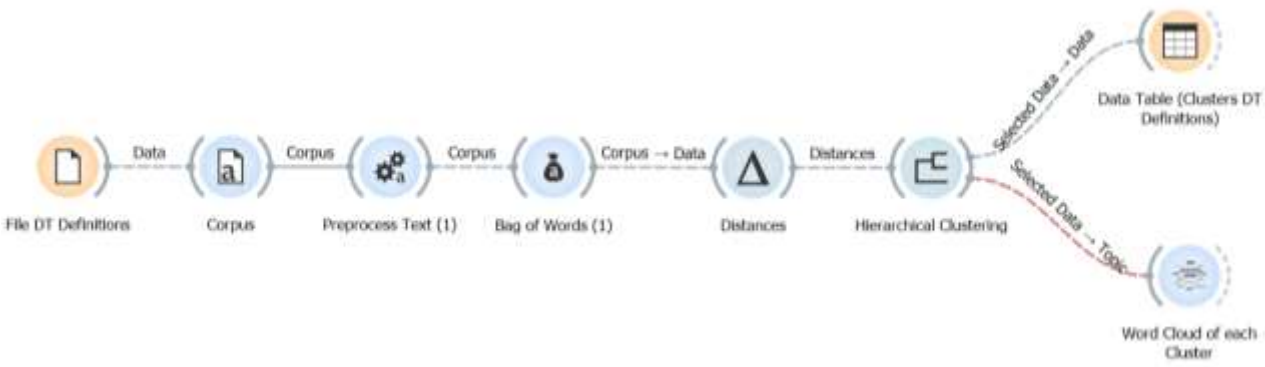


Figure 2: First model designed in Orange to automatically discover clusters in the DT definitions set

The second model has been designed for identifying which topics are the most debated and discussed in the selected DT papers. The statistical model that has been used is the topic modelling as shown in Figure 3. Topic modelling concerns using a text-mining tool for discovering hidden semantic structures in a text body. In our review, for each DT paper the authors name, title, abstract, keywords and content were collected, listed, and pre-processed for discovering the "topics" that occur in our selection employing the Latent Dirichlet allocation (LDA). Latent Dirichlet allocation is a statistical model that automatically detects a set of topic modelling, classifies papers, and estimate their relevance to various topics. The outcome is reported in *"Data Table (Topic Modelling)"*. A paper typically contains multiple topics in different proportions; thus, LDA also reports the topic weight per paper and this can be visualized in *"Distributions (Topic Modelling)"*. The results will be discussed in section 3 to design and compose a digital twin paradigm.

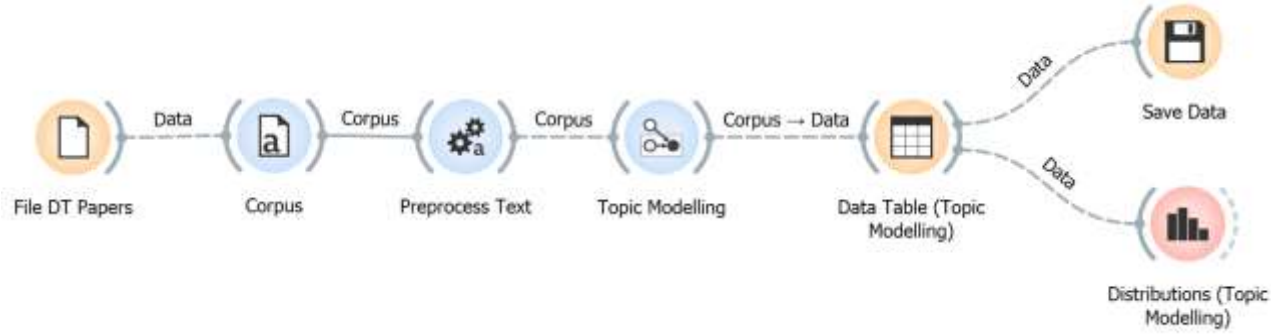


Figure 3: Second model designed in Orange for discovering Topic Modelling in DT papers

Step 3: Knowledge Representation: Formal Concept Analysis

A more detailed look through the topics discovered in step 2 was felt necessary to highlight the digital twin trends and strategies. This further analysis is conducted by Formal Concept Analysis (FCA). Formal Concept Analysis (FCA) is a mathematical theory oriented at applications in knowledge representation (Agrawal et al., 1993). It provides tools to group the data and to discover formal patterns by representing it as a hierarchy of formal concepts organised in a semi-ordered set named lattice (Wille, 2002). In formal concept analysis (FCA), a formal context is a triple $K = (O, A, R)$, where O and A are non-empty sets, and R is a binary relation between O and A ($R \subseteq O \times A$) (Ganter, Stumme, and Wille 2005). The formal context (O, A, I) of an input matrix of n rows and m columns

consists of a set of objects defined as O = $\{Obj_1, Obj_2, Obj_3, Obj_n\}$, a set of attributes defined as A=$\{Attr_1, Attr_{2,}\ Attr_3 \ldots Attr_m\}$ and a binary relation R defined as $Obj_i, Attr_j \in R$ if and only if the intersection of i-th row and j-th column is not blank (Škopljanac-Mačina and Blašković, 2014). The FCA data table is composed by the set of objects (O) in rows and the set of attributes (A) in columns as shown in Table 1. In our review, the objects are the papers selected in step 1, while the attributes are all the topics identified by the model presented in step 2, Figure 3. The symbol "•" denotes that there is a relationship (R) between the object and the attribute.

Table 1: FCA Table

O / *A*	$\boldsymbol{Attr_1}$	$\boldsymbol{Attr_2}$	$\boldsymbol{Attr_3}$	$\boldsymbol{Attr_m}$
$\boldsymbol{Obj_1}$	•		•	
$\boldsymbol{Obj_2}$		•		
$\boldsymbol{Obj_3}$			•	•
$\boldsymbol{Obj_n}$	•		•	•

Given a set of objects (O), a set of attributes (A), and defined the relations (R) between objects and attributes, a formal concept represents a subset of objects sharing the same subset of attributes, as displayed in Figure 4. Each node in the figure represents a concept. A concept is constituted by two parts: its extension which consists of all objects belonging to the concept, and its intention which comprises all attributes shared by those objects. This understanding allows the formal discovery of associations among concepts and consequently recognizing which concepts are closely related based on the set of shared attributes (Valtchev, Missaoui, and Godin 2004). The results provided by Formal concept analysis will be discussed in section 3 to explore the trends in the combination between the identified topics and their regularity of appearance in the literature.

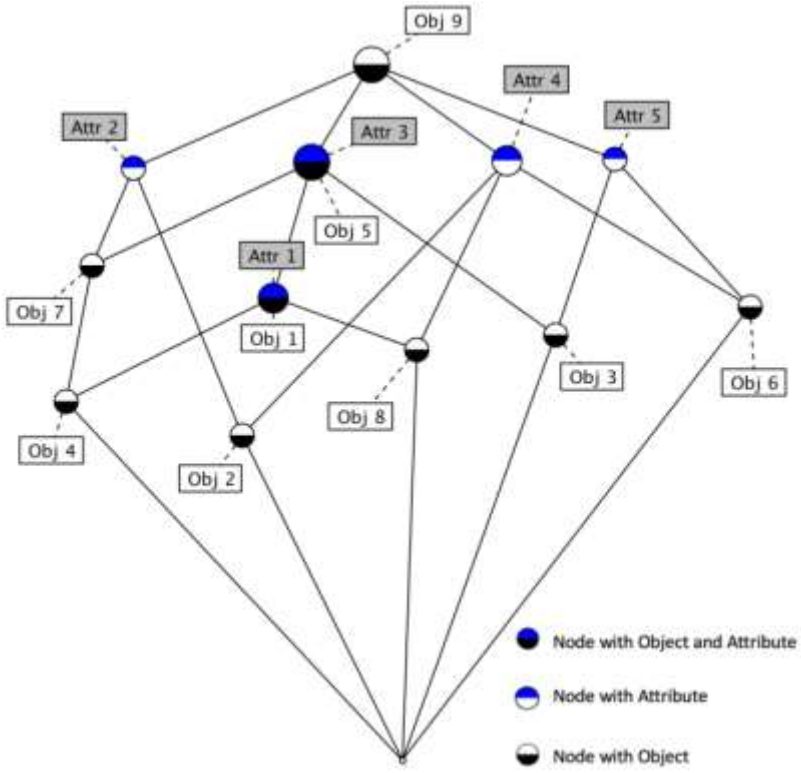


Figure 4: Example of Concept Lattice (Lezoche and Panetto, 2018)

2 State of Art Digital Twin: Definitions and Sights

Industry and academia define a digital twin in several different ways (Trauer et al., 2020). For example, according to some, a digital twin is a virtual representation/model that interacts with the physical system throughout its life cycle (Grieves and Vickers, 2017), (Glaessgen and Stargel, 2012). Other widely circulated definitions regard the need to exchange information between the two spaces involving sensors, data, and models (Lee et al., 2013), (Negri et al., 2017). Others consider a digital twin as the cyber part of a cyber-physical system (CPS) (Alam and El Saddik, 2017), (Graessler and Poehler, 2017). The concept of a digital twin has been investigated employing the model presented in Figure 2 to analyse the DT definitions listed below for understanding why it matters. The data table presents an additional column: *"Belonging Cluster"* that results from the application of the hierarchical cluster algorithm (HCA). This is a way to visualize how the DT definitions are grouped.

Table 2: Digital Twin Definitions

DIGITAL TWIN DEFINITIONS				
ID	**Year**	**Authors**	**DT Definition**	**Belonging Cluster**
1	**2002**	(Grieves,2014)	"a set of virtual information constructs that fully describes a potential or actual physical manufactured product from the micro atomic level to the macro geometrical level. At its optimum, any information that could be obtained from inspecting a physical manufactured product can be obtained	**C1**

			from its Digital Twin. The Digital Twin concept model contains three main parts: a) physical products in Real Space, b) virtual products in Virtual Space, and c) the connections of data and information that ties the virtual and real products together."	
2	**2012**	(Glaessgen and Stargel, 2012)	"an integrated multi-physics, multi-scale, probabilistic simulation of a complex product and uses the best available physical models, sensor updates, etc., to mirror the life of its corresponding twin."	**C1**
3	**2012**	(Tuegel, 2012)	"a cradle-to-grave model of an aircraft structure's ability to meet mission requirements, including sub-models of the electronics, the flight controls, the propulsion system, and other subsystems."	**C5**
4	**2013**	(Lee et al., 2013)	"a coupled model enables a digital twin of the real machine that operates in the cloud platform in parallel with the real process and simulates the health condition with an integrated knowledge from both data driven analytical algorithms as well as other available physical knowledge. The coupled model approach first constructs a digital image of a machine from the early design stage."	**C3**
5	**2015**	(Ríos et al., 2015)	"a product equivalent digital counterpart that exists along the product life cycle from conception and design to usage and servicing, knows the product past, current and possible future states, and facilitates the development of product related intelligent services."	**C1**
6	**2015**	(Rosen et al., 2015)	"a very realistic model of the current state of the process and their own behaviour in interaction with their environment in the real world."	**C4**
7	**2016**	(G. N. Schroeder et al., 2016)	"a virtual representation of the real product. It has product's information from the beginning of the life until the disposal of the product. The Digital Twin is a counter part of the physical device, machine or product in a CPS. It has the information related to the whole life cycle of a product."	**C1**
8	**2017**	(Alam and El Saddik, 2017)	"the cyber layer of CPS, which evolves independently and keeps close integration with the physical layer."	**C2**
9	**2017**	(Brenner and Hummel, 2017)	"a digital copy of a real factory, machine, worker etc., that is created and can be independently expanded, automatically updated as well as being globally available in real time."	**C3**
10	**2017**	(Ciavotta et al., 2017)	"a digital avatar encompassing CPS data and intelligence, representing structure, semantics, and behaviour of the associated CPS, and providing services to mesh the virtual and physical worlds."	**C4**
11	**2017**	(Graessler and Poehler, 2017)	"a cyber-physical device of its own, which is connected to the CPPS and tries to emulate the human employee through dynamically adapted values of a database, which represent for example properties, preferences, work schedule and skillset."	**C2**
12	**2017**	(H. Zhang et al., 2017)	"a set of realistic product and production process models linking enormous amounts of data to fast simulation and allowing the early and efficient assessment of the consequences, performance, quality of the design decisions on products and production line."	**C1**
13	**2017**	(Negri et al., 2017)	"a virtual and computerized counterpart of a physical system that can exploit a real-time synchronization of the sensed data coming from the field and is deeply linked with Industry 4.0."	**C5**
14	**2017**	(Schleich et al., 2017)	"a bi-directional relation between a physical artefact and the set of its virtual models, enabling the efficient execution of product design, manufacturing, servicing, and various other activities throughout the product life cycle."	**C1**
15	**2017**	(Schluse et al., 2017)	"a one-to-one virtual replica of a "technical asset" (e.g., machine, component, and part of the environment). A digital twin contains models of its data (geometry, structure, . . .), its functionality (data processing, behaviour, . . .), and its communication interfaces. It integrates all knowledge resulting from modelling activities in engineering (digital model) and from working data captured during real-world operation (digital shadow). A Digital Twin contains models of its "data" (geometry, structure, …), its functionality (data processing, behaviour, …) and its communication interfaces."	**C3**
16	**2017**	(Söderberg et al., 2017)	"a digital copy of a product or a production system, going across the design, pre-production, and production phases and performing real-time optimization."	**C1**
17	**2017**	(Stark et al., 2017)	"a unique instance of the universal Digital Master model of an asset, its individual Digital Shadow and an intelligent linkage (algorithm, simulation model, correlation, etc.) of the two elements above."	**C3**

18	**2017**	(Weber et al., 2017)	"a digital representation that contains all the states and functions of a physical asset and has the possibility to collaborate with other digital twins to achieve a holistic intelligence that allows for decentralized self-control."	**C3**
19	**2017**	(Yun et al., 2017)	"a perfect digital entity of a physical system; it accurately reflects the status of the corresponding physical machine. We can tightly control the system through a digital twin, that is, a cyber model of the machine."	**C3**
20	**2018**	(Autiosalo, 2018)	"the cyber part of a Cyber-Physical System."	**C2**
21	**2018**	(Asimov et al., 2018)	"a virtual replica of real physical installation, which can check the consistency for monitoring data, perform data mining to detect existing and forecast upcoming problems, and which uses an AI knowledge engine to support effective business decisions."	**C5**
22	**2018**	(Bao et al., 2018)	"a virtual model in the virtual space, and it is used to simulate the behaviour and characteristics of the corresponding physical object in real time."	**C4**
23	**2018**	(Lee and Kim, 2018)	"a near real-time digital image of a physical object or process that helps optimize business performance. Two concepts of IoT (Internet of things) and IoS (Internet of Service) are combined to realis e the smart factory based on a digital twin."	**C3**
24	**2018**	(Haag and Anderl, 2018)	"a comprehensive digital representation of an individual product. It includes the properties, condition, and behaviour of the real-life object through models and data. The digital twin is a set of realistic models that can simulate its actual behaviour in the deployed environment. The digital twin is developed alongside its physical twin and remains its virtual counterpart across the entire product life cycle."	**C3**
25	**2018**	(Luo et al., 2018)	"a complete virtual prototype of an entire system and a one-to-one mapping relationship. Therefore, a multi-domain digital modelling method is needed; a consistent model between the designed and the actual environment of a machine tool should be established, which needs the real-time and accurate data mapping method; an effective machine learning algorithm to mine the data gathered from sensors and the control system is also necessary."	**C5**
26	**2018**	(Nikolakis et al., 2018)	"a digital replica of the physical environment along with the operator. This model constrains the behaviour of the twin towards replicating the actions of the physical system's actuators."	**C4**
27	**2018**	(Tao et al., 2018b)	"a set of virtual models. These mirror images and mapping of the physical products in the virtual space. They could reflect the whole life cycle process, as well as simulate, monitor, diagnose, predict, and control the state and behaviours of the corresponding physical entities. The virtual models include not only the geometric models, but also all rules and behaviours, such as material properties, mechanical analysis, health monitoring."	**C4**
28	**2018**	(Z. Liu et al., 2018)	"a living model that continually adapts to change in the environment or operation using real-time sensory data and can forecast the future of the corresponding physical assets for predictive maintenance."	**C4**
29	**2018**	(Zhuang et al., 2018)	"a dynamic model in the virtual world that is fully consistent with its corresponding physical entity in the real world and can simulate its physical counterpart's characteristics, behaviour, life, and performance in a timely fashion."	**C4**
30	**2019**	(Leng et al., 2019)	"each physical device will have its cyber part as a digital representation of the real device, culminating in the digital twin models. So, the digital twin can monitor and control the physical entity, while the physical entity can send data to update and synchronize its virtual model."	**C3**

HCA detects five different clusters labelled: C1, C2, C3, C4, C5 as shown in Figure 5. To assist with the interpretation and verification of each cluster, word clouds were generated to provide additional evaluations identifying the occurrence of words shared by the grouped definitions.

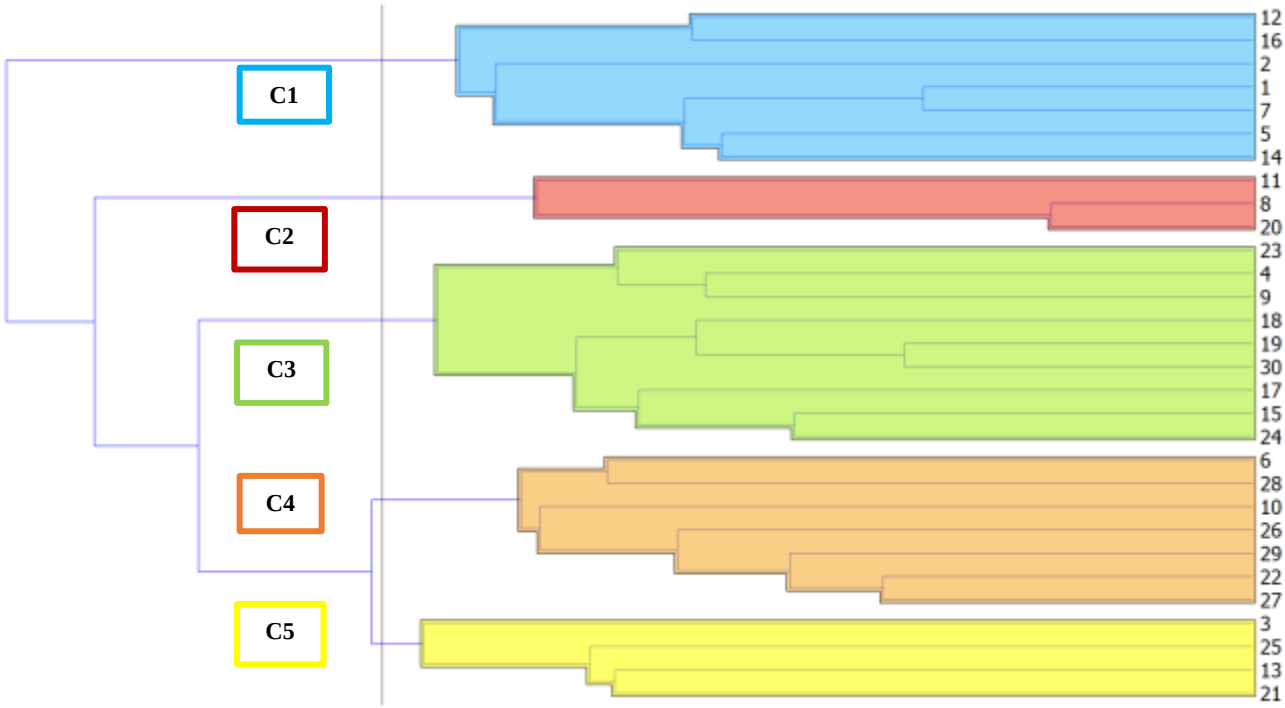


Figure 5: Hierarchical Clustering (HCA) Results - Clusters DT definitions

The cluster C1 involves the definitions provided by: (Grieves, 2014), (Glaessgen and Stargel, 2012), (Ríos et al., 2015), (G. N. Schroeder et al., 2016), (H. Zhang et al., 2017), (Schleich et al., 2017), (Söderberg et al., 2017). The corresponding word cloud in Figure 6 shows that the consideration of the life cycle phases is the core point in the definition of a Digital Twin. The concept of a Digital Twin was first mentioned in a presentation of the University of Michigan in 2002 entitled "Conceptual Ideal for PLM". As the concept was emerging out of the field of Product life cycle management (PLM), (Grieves and Vickers, 2017) referred to the connection between real space and virtual space over all phases of the product life cycle presenting all the elements of the Digital Twin: real space, virtual space, the link for data flow from real space to virtual space, the link for information flow from virtual space to real space and virtual sub-spaces. (Ríos et al., 2015) and (Schleich et al., 2017) specify the product life cycle from conception and design to usage and servicing while (G. N. Schroeder et al., 2016) suggest the existence of a Twin from the beginning of a product's life until its disposal. In (H. Zhang et al., 2017) the DT can integrate data in the product life cycle to accurately simulate and assess the performance and the quality of the design decisions on products and production lines. According to (Söderberg et al., 2017) a Digital Twin exists over the complete life cycle, subdivided in the phases design, pre-production, and production for performing real-time optimization. The basic idea behind a Digital Twin, in (Glaessgen and Stargel, 2012), is a high-fidelity virtual model of the physical entities having the scope of replicating and simulating the states and behaviours of these latter along its life.

The Digital Twin is defined as a new paradigm in simulation (Rosen et al., 2015). It extends the use of simulation to all phases of the product life cycle (Garetti et al., 2012), (Rodič, 2017). Simulation is the basis for design decisions, validation, and test not only for a generic device but also for

monitoring complete systems (Boschert and Rosen, 2016). The digital twin is the evolution of other research fields such as Virtual Manufacturing systems, Model-based Predictive control (MPC), and Building Information Modelling (BIM).

- Virtual manufacturing (VM) is defined as a system aimed at generating a virtual representation of a physical system without using real facilities/entities (Onosato and Iwata, 1993). In comparison with the characteristics of the Digital Twin, the key differences are the lack of connection between physical and virtual entities (Jones et al., 2020).
- The Model-based predictive control uses a model of the system to make predictions about the system's future behaviour (Ma et al., 2011). The digital twin and MPC simulates current state to change the future state but the DT aims to create virtual models in sync with their physical part.
- The Building Information Modelling (BIM) is a process for creating and managing a model containing digital information about a specific asset across its whole life cycle. The major differences between a building's BIM and a digital twin are that BIM is designed to improve the efficiency of design and construction and it does not work with real-time data. Whereas the Digital twin works specifically with real-time data to monitor a physical asset and improve its operational efficiency enabling the predictive manufacturing (Khajavi et al. 2019).

The DT is becoming increasingly relevant to Model-based systems engineering (MBSE) (Bachelor et al., 2019). MBSE is defined as: "*formalized application of modelling to support system requirements, design, analysis, verification, and validation activities beginning in the conceptual design phase and continuing throughout development and later life cycle phases*" (International Council on Systems Engineering (INCOSE), 2007). The simulation in MBSE is mostly considered to be a tool for R&D departments (Boschert and Rosen, 2016). The Digital Twin extends the concept of MBSE from engineering and manufacturing to the operation and service phases (Estefan, 2007), (Boschert et al., 2018). Digital twins could provide an effective implementation for realising the end-to-end integration of a system across the entire product life cycle (Cheng et al., 2020) because it is designed to interconnect and close the gap between product design, product development, production planning, production, and associated services (Weyer et al., 2016).

Figure 6: Word Cloud DT Definitions of Cluster C1

The real power of a digital twin is that it can provide a near-real-time comprehensive linkage between the physical and cyber systems. It is the core aspect reflected in the Word cloud in Figure 7, that

introduces the Cyber-Physical System (CPS) concept in DT definitions (cluster C2) provided by (Alam and El Saddik, 2017), (Autiosalo, 2018), (Graessler and Poehler, 2017). A Cyber-physical system (CPS) (Monostori et al., 2016) aims at embedding computing, communication and controlling capabilities (3C) into physical assets to converge the physical space with the virtual space. (Alam and El Saddik, 2017) and (Autiosalo, 2018) agree in defining the DT as the cyber part of a Cyber-Physical System to emulate the human employee through dynamically adapted values of a database, which represent for example properties, preferences, work schedule and skillset (Graessler and Poehler, 2017).

CPS and DTs involve the integration between the physical and the cyber space. However, DTs focus more on virtual models, while CPS emphasize 3C capabilities (computing, communication and controlling). In this way, the physical space can be integrated, monitored, controlled and coordinated by the virtual space in real-time and vice versa (Blume et al., 2014). Sensors and actuators are considered as key elements in CPS (Lee et al., 2015), while models and data are the core components in a DT (Tao et al., 2019). A digital representation, hence, a digital twin, bridges the gap between the physical and cyber system improving and supporting the decision making (Zhang et al., 2019).

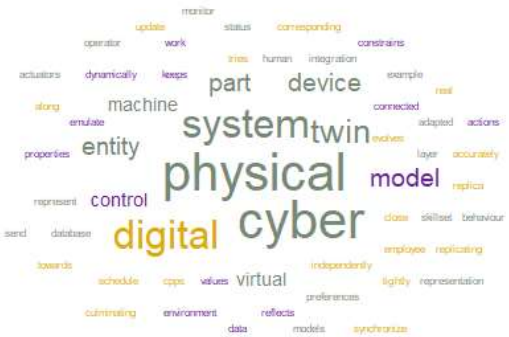


Figure 7: Word Cloud DT Definitions of Cluster C2

Some authors (Lee et al., 2013), (Brenner and Hummel, 2017), (Schluse et al., 2017), (Stark et al., 2017), (Weber et al., 2017) (Yun et al., 2017), (Lee and Kim, 2018), (Haag and Anderl, 2018), (Leng et al., 2019) involve a comprehensive representation of all data, information, and knowledge of the physical twin in the Digital Twin definitions. The word cloud in Figure 8 shows the results of definitions clustered in C3. Most definitions aiming at an increased convergence between real and virtual space. A closed loop needs to be realized, (Leng et al., 2018) between the physical and virtual worlds through real time data connection (Brenner and Hummel, 2017), (Schluse et al., 2017), (Stark et al., 2017) by collecting and analysing data (Weber et al., 2017), (Lee et al., 2013), (Haag and Anderl, 2018) to respond to the changes over the time and to optimize business performance (Lee and Kim, 2018). Even though the definitions agree on the virtual space, they differ regarding the physical space. Some refer to machines, or more generally to products (Lee et al., 2013), (Lee and Kim, 2018), (Haag and Anderl, 2018), (Leng et al., 2019). Others consider an asset as a physical space (Schluse et al., 2017), (Stark et al., 2017), (Weber et al., 2017). (Yun et al., 2017) and (Lee and Kim, 2018) extend the notion of physical space to the whole system or process. In more details the definitions converge to the descriptions of which type of data should be transferred from the physical to the

virtual space. (Schluse et al., 2017) consider not only data generated by the physical product, but also its functionality (data processing, behaviour) and its communication interfaces. (Haag and Anderl, 2018) go even further by defining the properties, condition, and behaviour of the real-life object. A Digital Twin is not just defined by the data. It also includes data driven analytical algorithms in (Lee et al., 2013) and in (Stark et al., 2017) to reflect the status of the corresponding physical part (Yun et al., 2017). A Digital Twin can integrate data from multiple sources. The interaction with the physical system should be bidirectional (Leng et al., 2019). Data collected from the physical space updates the virtual model. The physical twin improves its performance during real time operation exploiting knowledge acquired from the data.

Figure 8: Word Cloud Definitions of Cluster C3

Regardless of the represented physical space, it needs to be defined which aspects of the physical space should be transferred to the virtual space. The disunity in literature on how to model the behaviour of the physical space is even present in the definitions belonging to cluster C4 (Rosen et al., 2015), (Ciavotta et al., 2017), (Bao et al., 2018), (Nikolakis et al., 2018), (Tao et al., 2018b), (Z. Liu et al., 2018), (Zhuang et al., 2018), shown in Figure 9. Some try copying the physical behaviour, its properties, and characteristics in very realistic (Rosen et al., 2015) virtual models to simulate the behaviour of the current status of the physical space (Bao et al., 2018). The need for a set of virtual models stems from the fact that the virtual models include not only the geometric models, but also all rules and behaviours, such as material properties, mechanical analysis, health monitoring (Tao et al., 2018b) to monitor, diagnose, predict, and control the state and behaviours of the corresponding physical entities (Tao et al., 2018b), (Nikolakis et al., 2018). Data and information should also consider all perspectives of the physical space including, structure, semantics, and behaviour to mesh the virtual and physical worlds (Ciavotta et al., 2017). The DT is typically applied in contexts characterized by uncertainty and complexity, where the working conditions may vary depending on external and internal factors. For this reason, (Z. Liu et al., 2018) propose the concept of 'living model' while (Zhuang et al., 2018) the concept of 'dynamic model' i.e., a model that continually adapts and changes in the environment. The Digital Twin should evolve synchronously with the real system along its whole life cycle. It should be able to modify its initial configuration and to adapt itself to the current situation. This aspect introduces another feature debated in the literature, namely

the difference between the simulation capabilities and the emulation capabilities of a DT. On the one hand, the simulation capabilities of a DT are provided by a design of its environment allowing to approximate the behaviour of the real systems to represent how the system reacts (Law et al., 2000). It can be thought of as a “static feature” of the DT. On the other hand, the emulation refers to the capability of a DT to be synchronous with the real system, so as it behaves almost similarly to the actual behaviour of the physical system (Ayani et al., 2018). Accordingly, this feature of DT can be thought of as a “dynamic feature”. An emulation model operates in a hardware-in-the-loop configuration to perform the same work of the physical system. It provides a closer replication with respect to the simulation model (Lee and Park, 2014). From the simulation point of view, the digital twin represents a new wave in modelling and simulation (Rosen et al., 2015). From the emulation point of view, the digital twin duplicates and imitates the physical system in the virtual word. It can thus help to proactively understand what should be done and to react to modifications in the real world.

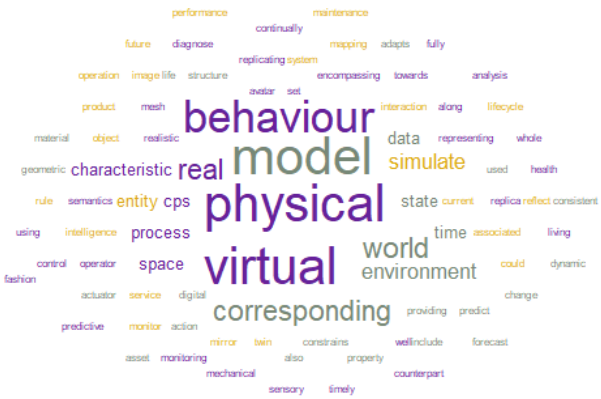


Figure 9: Word Cloud DT Definitions of Cluster C4

The virtual system concept, in Figure 10, sums up the DT definitions clustered in C5 provided by (Tuegel, 2012), (Negri et al., 2017), (Asimov et al., 2018), (Luo et al., 2018). The virtual system enables the replication of the physical system into its “digital twin” throughout the entire value chain, by merging data into behaviour models (Borangiu et al., 2019). The physical twin automatically transfers data of its behaviour, its status, and information to the virtual space over the entire life cycle. The virtual system defined also as virtual replica in (Asimov et al., 2018), virtual prototype in (Luo et al., 2018), and virtual counterpart in (Negri et al., 2017), provides different services such as the control of the current situation and the prediction of the near future and sends them back to the real space so the physical product or process adapt accordingly.

The virtual system may enable companies and organisation to solve physical issues faster by detecting them sooner, predict outcomes, design, and build better products, and ultimately, better serve their customers (Trauer et al., 2020).

Figure 10: Word Cloud DT Definitions of Cluster C5

The ability of simulation along product life cycle (C1), the synchronization of the cyber system with the physical assets (C2), the integration of real time data (C3), the behavioural modelling of the physical space (C4), the services provided by the virtual system (C5) are the main aspects that characterize the digital twin definition. In view of above, what it is a digital twin can be summarized as follow:

"*A set of **adaptive models** that emulate the **behaviour** of a physical system in a **virtual system** getting **real time data** to update itself along its **life cycle**. The digital twin replicates the physical system to predict failures and opportunities for changing, to prescribe real time actions **for optimizing and/or mitigating** unexpected events observing and evaluating the operating profile system".*

3 Digital Twin Paradigm

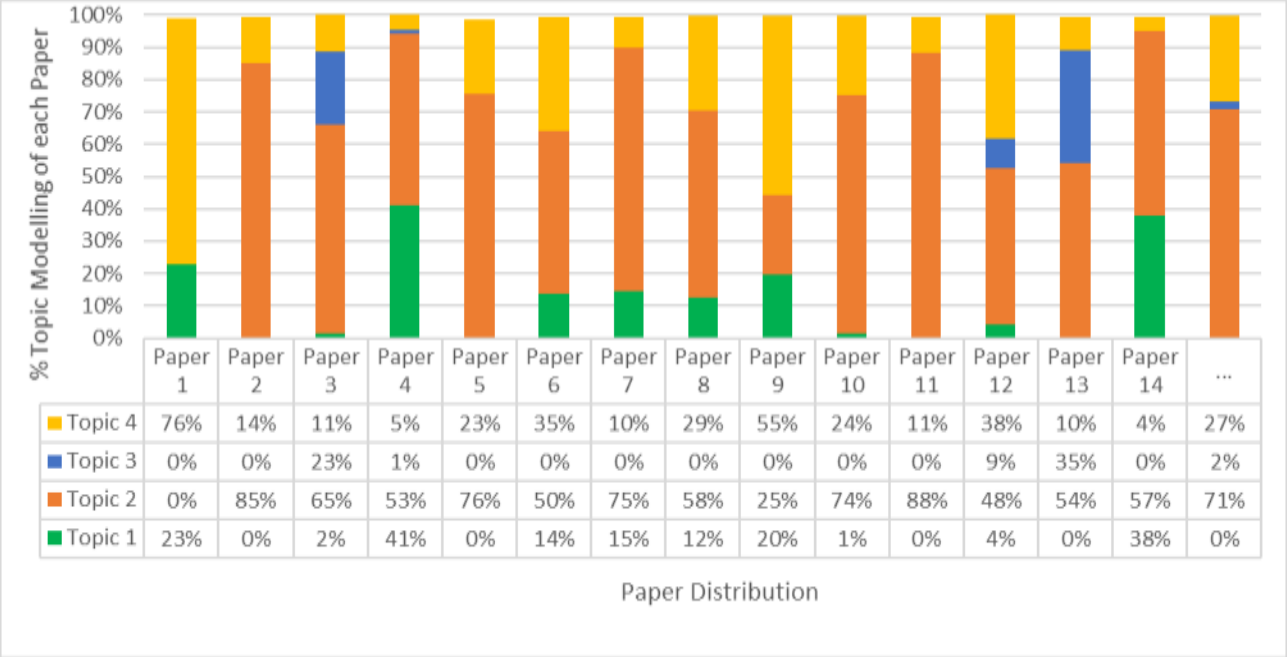
Existing scientific research on DT's, are focussed on different application domains and their related technologies. A more detailed look is needed to define where, when, why and how develop a digital twin. The model presented in Figure 3 and explained in section 1 (step 2), has been applied to automatically discover the main topics in the DT collection and classify them into the discovered topics. Four different topic modelling have been detected applying the Latent Dirichlet Allocation (LDA) algorithm. The results can be visualized in table 3 that presents the set of keywords that make up each topic modelling. On the base of the concepts introduced and analysed also in the state of the art it is possible to define the semantic of the topics below. The first topic modelling presents the keywords: *'support-production, decision-making, simulation, analysis'* indicating that different studies in literature review have been devoted to the analysis of the digital twin functions. The analysis of papers classified in the first topic modelling allows to evaluate why the digital twin should be used. The second topic modelling introduces and considers the enabling components for designing and developing a Digital Twin as: '*data-driven and physical model'* allowing to evaluate which are the key components and technologies employed for implementing digital twins. The words of the third topic modelling reveal that the value of the digital twin is its ability to understand and to predict how the product will perform in production, and to optimize its services and performance through its life cycle. It means that the set of papers belonging to topic 3 treat where the digital twin should be

applied and when should be developed. The papers grouped in the fourth topic modelling are focused on the exploration of the digital twin *configuration/architecture*. The *architecture* is the basic principle to investigate for understanding how to design a digital twin.

Table 3: Data Table (Topic Modelling) provided by the model designed in Orange (Figure 3)

ID	TOPIC MODELLING: Set of Words	TOPIC LABEL
TOPIC 1 (T1)	**Support-production, decision-making, simulation, analysis,** approach	**FUNCTIONS** *'Why should a Digital Twin be used?'*
TOPIC 2 (T2)	**Big data, data-driven,** management, **physical model,** shopfloor, system	**COMPONENTS/ TECHNOLOGIES** *'How to implement a Digital Twin?*
TOPIC 3 (T3)	**Life cycle**, cps, level, **product, process, service, improve,** application	**CONTEXT and LIFE CYCLE** *'Where is appropriate to use a Digital Twin?'* *'When has a Digital Twin to be developed?'*
TOPIC 4 (T4)	**Architecture**, **virtual framework,** smart manufacturing	**ARCHITECTURE** *'How to design a Digital Twin?'*

A paper typically can cover multiple topics in different proportions (%). LDA algorithm automatically classifies the DT papers to topics and estimates their relevance to each topic as shown in Figure 11. For example, (Abramovici et al., 2017) (Paper 1) covers the topics 1 and 4 at 23% and 76% respectively.



	Paper 1	Paper 2	Paper 3	Paper 4	Paper 5	Paper 6	Paper 7	Paper 8	Paper 9	Paper 10	Paper 11	Paper 12	Paper 13	Paper 14	...
Topic 4	76%	14%	11%	5%	23%	35%	10%	29%	55%	24%	11%	38%	10%	4%	27%
Topic 3	0%	0%	23%	1%	0%	0%	0%	0%	0%	0%	0%	9%	35%	0%	2%
Topic 2	0%	85%	65%	53%	76%	50%	75%	58%	25%	74%	88%	48%	54%	57%	71%
Topic 1	23%	0%	2%	41%	0%	14%	15%	12%	20%	1%	0%	4%	0%	38%	0%

Figure 11: Topic Modelling Distribution of each Paper

The overview distribution of all topics is provided in Figure 12. It illustrates that 46,7% of papers covers the DT functions, DT components and DT architectures (T1ΛT2ΛT4) topics while 31,4% of papers analyse all topics. All the absent combinations are with percentage of zero.

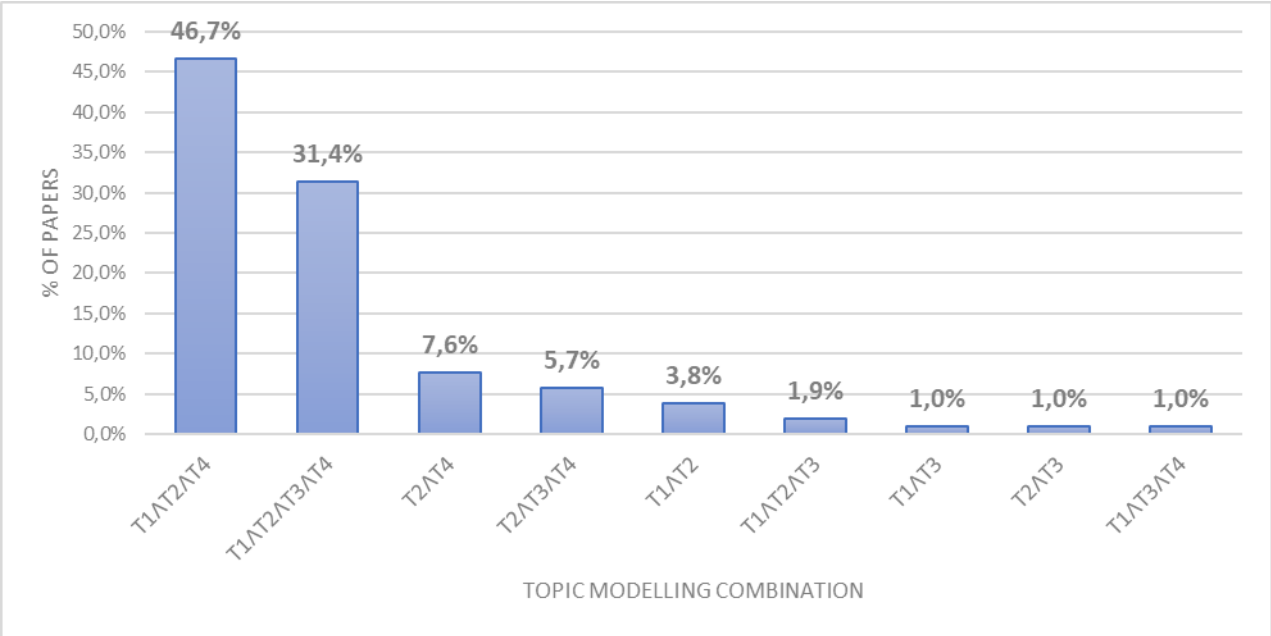


Figure 12: Overview Distribution of Topic Modelling

The topic modelling analysis distribution allows understanding that several studies have been devoted to analysing the Digital Twin concept and its instantiation in different application **contexts (T3)**. At the same time, in each context, the digital twins have their own specificity as functions in the **life**

cycle (T3) phases namely design, manufacturing and service. As a result, each application of a DT varies depending on a different **function (T1)** accordingly. However, the design of a digital twin requires the definition of an **architecture (T4)** and the enabling **components/technologies (T2)** to implement it. As a summary of the topic modelling performed above, these can be logically sorted to shape the digital twin paradigm as follow:

1. Application Contexts (Where is appropriate to use a Digital Twin?), TOPIC MODELLING 3
2. Life cycle (When has a Digital Twin to be developed?), TOPIC MODELLING 3
3. Functions (Why should a Digital Twin be used?), TOPIC MODELLING 1
4. Architecture (How to design a Digital Twin?), TOPIC MODELLING 4
5. Components/Technologies (How to implement a Digital Twin?), TOPIC MODELLING 2

The papers thus classified by LDA were analysed to identify and define which and how many subtopics each topic consists of. By subtopic we mean the identification of the main categories that characterise each topic. This allows us to compare and review the existing works to answer the main questions posed above. Table 4 shows the results. It reports one paper on each row and the topics and their corresponding sub-topics on columns. The *"Context and Application"* columns describe the application context taken into consideration in each article. Accordingly, to the papers belonging to this topic, the application contexts can be grouped in five categories, that are listed as follows: *Healthcare; Maritime and Shipping; Manufacturing; City Management; Aerospace*. The column under *"life cycle"* is split into the main product life cycle phases namely: *design, production, and service.* The *"Functions"* column defines the DT functions/purposes specifically: *Accelerating the product development speed; Identifying customers' needs; Performance optimization and validation; Remote commissioning and diagnostics*. The columns under *"Architecture"* reports the main layers used to design a DT namely: *Physical; Network; Computing*. The analysis also considers the *"Components"* i.e., the most discussed and applied technologies for building a Digital Twin. The same table was used for the analysis conducted by Formal Concept Analysis (FCA) to explore the trends and the combinations in literature on the design and the development of a digital twin. The results are discussed in the next sections presenting respectively the DT application contexts in section 3.1, the DT life cycle and its functions in 3.2, and the DT architecture and its components/technologies in 3.3.

Table 4: FCA Table Structure: Objects and Attributes

			REFERENCES DT PAPERS	1	2	3
T3	T3.1: CONTEXT AND APPLICATION	A1	DIGITAL TWIN IN AEROSPACE			
		A2	DIGITAL TWIN IN MARITIME AND SHIPPING			
		A3	DIGITAL TWIN IN MANUFACTURING	•	•	•
		A4	DIGITAL TWIN IN CITY MANAGEMENT			
		A5	DIGITAL TWIN IN HEATHCARE			
	T3.2: LIFE-CYCLE	A6	DIGITAL TWIN IN THE DESIGN PHASE			
		A7	DIGITAL TWIN IN THE PRODUCTION PHASE	•	•	
		A8	DIGITAL TWIN IN THE SERVICE PHASE			
T1	T1: FUNCTIONS	A9	ACCELERATING THE PRODUCT DEVELOPMENT SPEED			
		A10	IDENTIFYING CUSTOMERS' NEEDS			
		A11	PERFORMANCE OPTIMIZATION AND VALIDATION	•	•	
		A12	REMOTE COMMISSIONING AND DIAGNOSTICS			
T4	T4: ARCHITECTURE	A13	PHYSYCAL LAYER	•		
		A14	NETWORK LAYER	•		
		A15	COMPUTING LAYER	•	•	
T2	T2: COMPONENTS	A16	RFID	•		
		A17	WIRELESS SENSOR NETWORKS (WSN)	•		
		A18	RFID SENSOR NETWORKS			
		A19	UNIT LEVEL		•	
		A20	SYSTEM LEVEL	•	•	
		A21	SYSTEM OF SYSTEMS LEVEL			
		A22	MIDDLEWARE: SERVICE ORIENTED ARCHITECTURE (SOA)			
		A23	COMMUNICATION PROTOCOL	•		
		A24	COMMUNICATION PROTOCOL/INTERFACE: AUTOMATION ML	•		
		A25	WIRELESS COMMUNICATION			
		A26	PROGRAMMING INTERFACES (API)			
		A27	DATA-DRIVEN METHODS		•	
		A28	GEOMETRY MODEL	•		
		A29	PHYSICAL MODEL	•		•
		A30	BEHAVIOUR MODEL	•		•
		A31	COLLABORATIVE INFORMATION MODEL			
		A32	DECISION-MAKING/ RULE MODEL	•		
		A33	SCALABILITY;			
		A34	MODEL INTEROPERABILITY (e.g FMU)			
		A35	FIDELITY			
		A36	DINAMICITY			
		A37	MODULARITY			
		A38	INTERFACE LAYER (Web Services, Augmented reality)	•		

3.1 Digital Twin Application Contexts ('Where is appropriate to use a Digital Twin?' and 'Who is doing Digital Twins?')

The main papers covering the topic modelling 3 are discussed in this subsection to define and discuss the main DT application contexts. The review reveals that the most explored contexts are: Healthcare; Maritime and Shipping; Manufacturing; City Management; Aerospace. Digital Twins in Healthcare is used for capturing and visualizing a hospital system to create a safe environment and to test the impact of potential changes on system performances. A digital twin can be used to predict the

outcome of specific procedures. It can determine the better therapy option for a specific patient. In healthcare, a digital twin recording data of a person, combined with AI models, can provide answers for clinical problems (Bruynseels et al., 2018). Digital Twins in Maritime and Shipping are used as support for design. The design requires to invest significant amounts of time and money in preparing analytical models to perform simulations. The digital twin allows to visualize all key components, to perform analyses and calculations, and to improve the control of the effects of operation on the ship's structural and functional components (Arrichiello and Gualeni, 2019).

A Digital Twin in Manufacturing involves different applications based on the stages across the entire lifetime of a product, such as design, production, logistics and maintenance (Dassisti et al., 2017), (Greif et al., 2020). The digital twin can support decision makers to predict an upcoming equipment failure, to inform an operator when an asset begins to show signs of non-optimal performance, to improve customer experience (Tao et al., 2018a).

Cities are areas of human settlement, with high population density, complete infrastructure, and buildings. Digital Twins, in City Management, improve the urban environment and people's quality of life. The digital twin can simulate people movements and emergency evacuations, modelling smart buildings, road traffic, air quality, infrastructure, and circular urban economies. The benefits of modelling range from preventive maintenance to operational efficiencies and cost savings. The DTs improve services for citizens, and increase safety and security (Mohammadi and Taylor, 2017). Aerospace companies have begun utilizing digital twins to accomplish the goal of reducing unplanned downtime for engines and other systems. Digital Twins in Aerospace may allow receiving advance warnings and predictions, but also preparing a plan of actions based on simulated scenarios that consider the weather conditions, the performance of the asset, and several other variables (Tuegel et al., 2011). With the help of digital twins, it is possible to develop and implement predictive maintenance for increasing the platform's operational availability and efficiency, extending its life cycle and reducing its cost. Moreover, DTs are capable of mitigating damage or degradation by activating self-healing mechanisms or by recommending changes in the mission profile to decrease loadings (Mandolla et al., 2019).

Digital Twins have attracted strong interests from industries too: GE Predix Platform, SIEMENS PLM, Microsoft Azure, IBM Watson, PTC Thing Worx, Aveva, SAP Leonardo Platform, Twin Thread, DNV-GL, Dassault 3D Experience, Sight Machine, Oracle Cloud. Patents have been filed by (Hershey et al., 2017) for General Electric and by (Song and Canedo, 2016) and (Fischer and Heintel, 2017) for Siemens. The Figure 13 shows the main DT platform for each application context. The roles of digital twins along life cycle management and its functions are discussed in the next section.

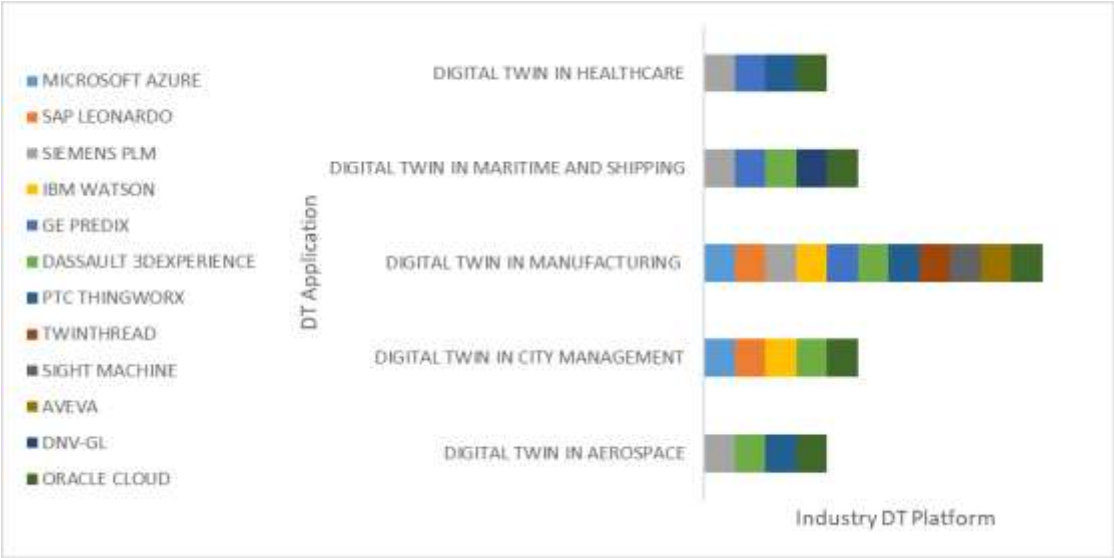


Figure 13: Industry DT Platform in each type of DT Application

3.2 Digital Twin Life cycle ('When has a Digital Twin to be developed?) and its Functions ('Why should a Digital Twin be used?')

In principle, out of the literature review and the DT definitions analysis presented in section 2, the digital twin finds application in the entire product life cycle management (PLM) that can be divided into three phases: Design; Production; Service (Bao et al., 2018), (Tao et al., 2018a). Regardless of the context domain, the DT has a series of functions in each phase of life cycle (Barricelli et al., 2019) that can be summarized in: Accelerating the product development speed; Identifying customers' needs; Performance optimization and validation; Remote commissioning and diagnostics. The existing trends and associations between DT life cycle and its functions was carried out by using Formal Concept Analysis (FCA). The lattice, illustrated in Figure 14, represents the hierarchy of concepts that group the papers according to their common life cycle phase and/or functions.

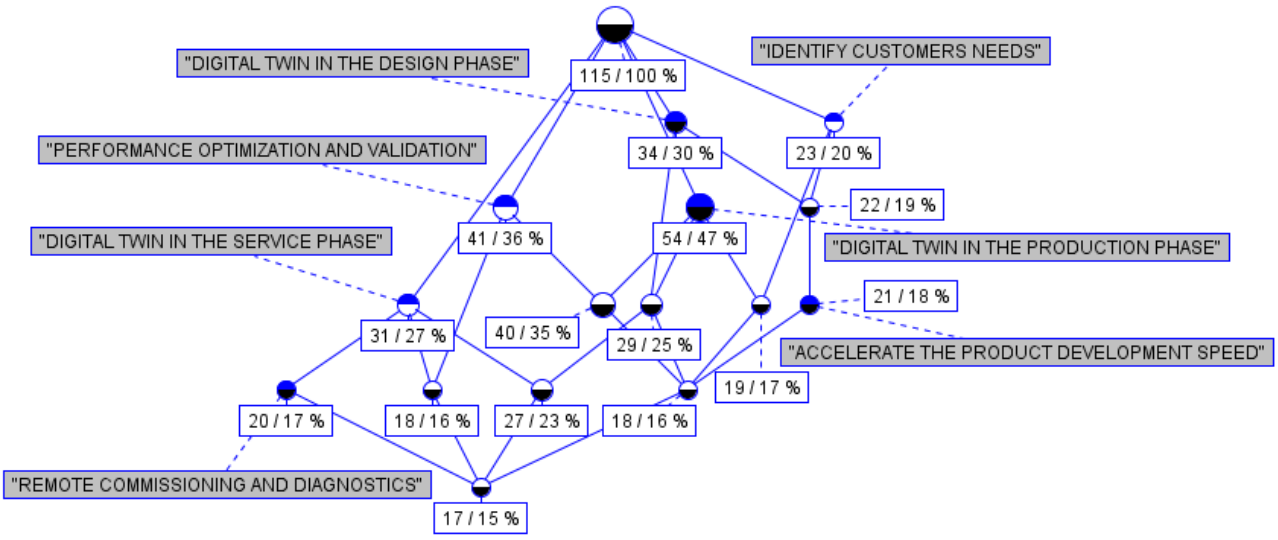


Figure 14: Lattice of Life Cycle phases and Functions

From Table 5, we can deduct that the FCA graph detects 16 different concepts. As explained in section 1 (step 3), a concept is constituted by two parts: its extension which consists of all objects belonging to the concept, and its intention which comprises all attributes shared by those objects.

Table 5: Formal Concepts of Life Cycle phases and Functions

FORMAL CONCEPTS	
ID CONCEPT	<{CONCEPT EXTENTS},{CONCEPT INTENTS}>
C1	<{115 Papers}>
C2	<{34 Papers}, {Digital Twin in the design phase}>
C3	<{23 Papers}, {Identify customers' needs}>
C4	<{41 Papers}, {Performance optimization and validation}>
C5	<{54 Papers}, {Digital Twin in the production phase}>
C6	<{22 Papers}, { Digital Twin in the design phase, Identify customers' needs}>
C7	<{31 Papers}, {Digital Twin in the service phase}>
C8	<{40 Papers}, {Digital Twin in the production phase, Performance optimization and validation}>
C9	<{29 Papers}, {Digital Twin in the design phase, Digital Twin in the production phase}>
C10	<{19 Papers}, {Digital Twin in the production phase, Identify customers' needs}>
C11	<{21 Papers}, {Digital Twin in the design phase, Identify customer's needs, Accelerate the product development speed}>
C12	<{20 Papers}, {Digital Twin in the service phase, Remote commissioning and diagnostics}>
C13	<{18 Papers}, {Digital Twin in the service phase, Performance optimization and validation}>

C14	<{27 Papers}, {Digital Twin in the design phase, Digital Twin in the production phase, Digital Twin in the service phase}>
C15	<{18 Papers}, {Digital Twin in the design phase, Digital Twin in the production phase, Digital Twin in the service phase, Identify customer's needs, Performance optimization and validation}>
C16	<{17 Papers}, {Digital Twin in the design phase, Digital Twin in the production phase, Digital Twin in the service phase, Identify customer's needs, Accelerate the product development speed, Performance optimization and validation, Remote commissioning and diagnostics}>

The concept C2 shows, through its Extent column, the existence, in our literature review, of {34 papers} which analyse the application of the <{Digital Twin in the design phase}>. The digital twin in the design phase can be applied to the conceptual design, detailed design, and virtual verification (Tao et al., 2018a) of a product. The digital twin in design stage is designed to generate the digital product design before the real execution (Q. Liu et al., 2018), (H. Zhang et al., 2017). In the conceptual design stage, the digital twin serves to guide designers to formulate functional requirements (Tao et al., 2018b). It can make the communication between customers and designers more transparent and faster by using the real-time transmission data (Tao et al., 2018a). In the detailed design phase, the digital twin enables simulation tests to ensure that the prototype can achieve the desired performance (Wärmefjord et al., 2017). In the virtual verification phase, the digital twin enables to simulate and predict the performance of the physical products based on virtual models (Damiani et al., 2018), (Bohlin et al., 2017). The concept C6 demonstrates that {22 Papers} in our selection, analyse the relation between the <{Digital Twin in the design phase} with the function {Identify customers' needs}>. Performances, customer usages and preferences are reflected in the twin, and then feed into the product development process to increase the customer satisfaction and market share (Tao et al., 2018b), (Macchi et al., 2018). The studies {21 Papers} grouped in C11 discuss the application of the < {Digital Twin in the design phase} for two different functions {Identify customer's needs, Accelerate the product development speed}>. Digital twins in the design phase can guide the designers to iteratively adjust the customers' expectations and improve the design models, achieving personalized design (Tao et al., 2018a). The digital twin can be used for designing products, testing them in real time situations, stipulating how the customer or the end user will use them and how the design will complement the product's environment (Söderberg et al., 2017). Data from the real machine are loaded into the digital model to enable simulation and testing of ideas even before actual manufacturing starts. The digital twin can be used to plan, reconfigure the product in response to external changes.

The concept C5 shows the existence of {54 papers} which analyse the application of the <{Digital Twin in the production phase}>. For example, (Leng et al., 2018) presents a Digital Twin for manufacturing cyber-physical systems (MCPS). (Ding et al., 2019) introduces a DT-based Cyber-Physical Production System (DT-CPPS). MCPS is used for controlling the shop floor manufacturing while DT-CPPS for improving the flexibility, controllability, and efficiency of shop floor manufacturing. A digital twin for production control and optimization can analyse the online data collected from the physical line for searching the optimal solution to the physical line (Sun et al., 2017) or to complex product assembly shopfloors (Zhuang et al., 2018). It can evaluate autonomously

the production real-time (Vachálek et al., 2017) and optimize the resource allocation (H. Zhang et al., 2018) autonomously (Rosen et al., 2015). A Digital Twin reference model for rotating machinery fault diagnosis was developed in (Wang et al., 2018), defining the requirements for constructing the Digital Twin model. A digital twin for hydraulic supports (Xie et al., 2019) is built to simulate the actual hydraulic and to support diagnosis and degradation analysis. The digital twin finds application also in CNC machine tool (Luo et al., 2018) and in smart injection process (Liau et al., 2018) to control the behaviours of the physical system in real-time. The papers {40 Papers} in C8 treat the application of <{Digital Twin in the production phase} for {Performance optimization and validation}>. Digital twins in the production phase aim at real time monitoring and optimization and for predicting the future state of the physical twin, thus preventing downtime and failures (Lee et al., 2013). The digital twin helps at determining the optimal set of parameters and actions that can help maximizing some of the key performance, and providing forecasts for long-term planning (Vachálek et al., 2017). The digital twin can analyse performance data collected over time and under different conditions (Alcácer and Cruz-Machado, 2019), reducing unplanned machine downtime, the amount of 'scrap' produced in each production line, and minimizing costly production quality faults. The DT can optimize and elevate the production process to a higher level of effectiveness and flexibility (Cimino et al., 2019)**.**

The concept C7 indicates that {31 papers} are oriented towards the <{Digital Twin in the service phase}>. The service phase refers to the phases after sale, including the product utilization and the maintenance (Tao et al., 2018a). (Abramovici et al., 2017) introduce a cloud-based Smart Product platform for the reconfiguration of Smart Products during the use phase using the concept of virtual product twins and an Internet of Things. The conceptual approach is prototypically demonstrated by considering a model environment for smart cars, which are temporarily reconfigured during their use phase. The digital twin has been developed also for the waste electrical and electronic equipment recovery to support the manufacturing/remanufacturing operations (Wang and Wang, 2018). The <{Digital Twin in the service phase} supports the {Remote commissioning and diagnostics}> of the operations of interconnected systems such as manufacturing systems, as presented in the studies {20 Papers} grouped in concept C12. This allows virtual monitoring systems and validation of the current status of production systems (i.e., energy monitoring and fault monitoring) (Qi et al., 2018). In addition, {18 Papers} analyse the application of the {Digital Twin in the service phase} for <{Performance optimization and validation}>. The digital twin can upgrade personalized product functions (Cheng et al., 2020) by obtaining the user's usage. In fact, in the service phase, Digital Twins can provide value-added services support for the prognostics and health management (PHM) (Qi et al., 2018) (Wang et al., 2018). The PHM is an engineering process of failure prevention and predicting reliability and remaining useful lifetime (RUL) (Sutharssan et al., 2015). In this case, the digital twin (DT) is developed for improving the accuracy and efficiency in the life cycle monitoring of a product (Tao et al., 2018c), (M. Zhang et al., 2018). There are currently relatively few digital twin applications {17 papers} for supporting the entire product life cycle (C16).

3.3 Digital Twin Architecture and Components (‘How to design and implement a Digital Twin?’)

A general and standard architecture of a digital twin was first built by (Grieves, 2014) that presents a physical space, a virtual space and the connection between them. There are various understandings of the DT architectures among researchers (Dassisti et al., 2017). (Stark et al., 2017) characterizes the DT as (1) an unique instance of the universal Digital Master model of an asset, (2) its individual Digital Shadow and (3) an intelligent linkage (algorithm, simulation model, correlation, etc.) of the two elements above (Kritzinger et al., 2018). An extended five-layer DT is proposed by (Tao et al., 2018c) and it is composed by: (1) Physical entity (PE); (2) Virtual entity (VE); (3) Services (Ss) for PE and VE; (4) Data (DD); (5) Connection (CN) among PE, VE, Ss and DD. Compared to Grieves’s architecture, data and services layers were added. The five-layer DT architecture developed by (Ponomarev et al., 2017) presents: (1) cyber-physical layer; (2) primary processing/store data layer; (3) distributed computing and storage layer; (4) models and algorithms layer; (5) visualisation and user interfaces layer. This kind of architecture highlights the data storage, the distributed computing and management system as critical parts of the digital twin. An extended six-layer DT is presented by (Redelinghuys et al., 2019). The layers are: (1) physical devices; (2) local controllers; (3) local data repositories; (4) IoT gateway; (5) cloud-Based information repositories; (6) emulation and simulation. This structure is more focused on the transmission of data flow from the physical system (Layer 1 e 2) to the cloud (Layer 5). From the computational perspective, the key functionality of a digital twin is the combination of physics-based models and data driven models to emulate and simulate the physical space accurately (Kaur et al., 2020).

In view of above, a DT architecture can be thought as consisting of several components and technologies organised into three main layers: the physical layer; the network layer; the computing layer (Boje et al., 2020). The physical layer consists of physical entities identified based on the stage of the product life cycle. The network layer connects the physical domain to the virtual one. It shares data and information. The computing layer involves the virtual entities emulating the corresponding real entities, including data-driven models and analytics, physic-based models, services, and users. Each layer is characterised by some DT components (for example hardware or software technologies, models, information structures) with commonalities in their scope of use and interactions, having also complementary functionalities. FCA was run on to detect which are the main and the most studied components/technologies for each layer. The formal concepts of the physical layer are shown in Figure 15 and in Table 6.

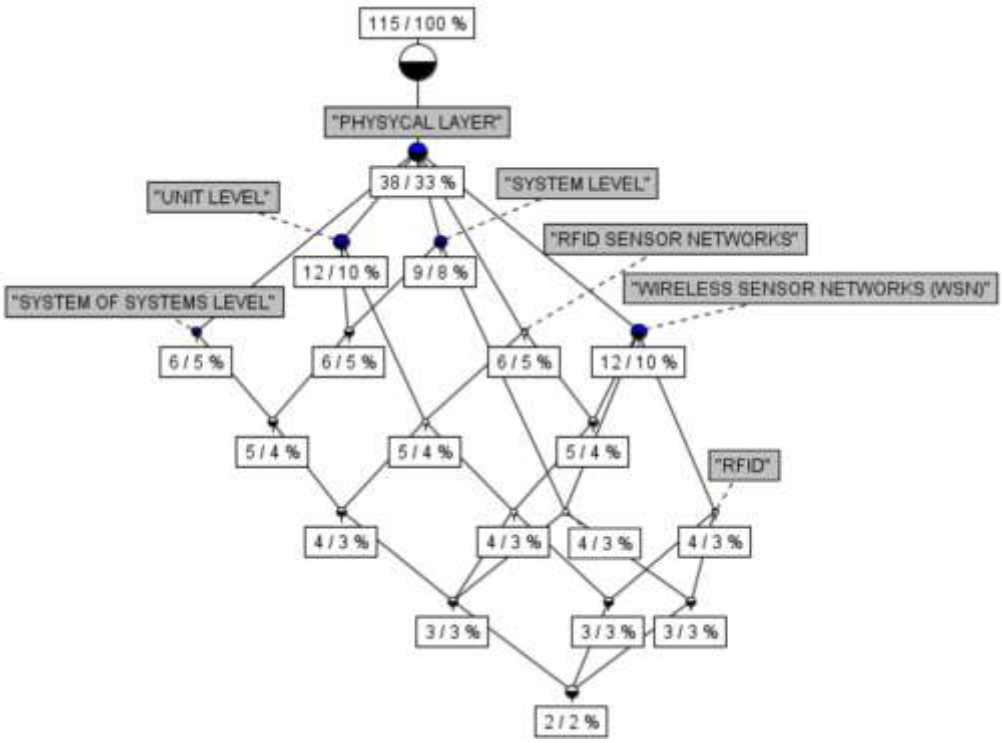


Figure 15: Lattice of Physical layer and its Components/Technologies

Table 6: Formal Concepts of Physical layer and its Components/Technologies

FORMAL CONCEPTS	
ID CONCEPT	<{CONCEPT EXTENTS},{CONCEPT INTENTS}>
C1	<{115 Papers}>
C2	<{38 Papers}, {Physical Layer}>
C3	<{12 Papers}, {Physical Layer, Unit Level}>
C4	<{9 Papers}, {Physical Layer, System level}>
C5	<{6 Papers}, {Physical Layer, System-of-systems level}>
C6	<{6 Papers}, {Physical Layer, Unit Level, System level }>
C7	<{6 Papers}, {Physical Layer, RFID sensor networks}>
C8	<{12 Papers}, {Physical Layer, Wireless sensor networks (WSN), RFID}>
C9	<{5 Papers}, {Physical Layer, Unit Level, System level, System-of-systems level}>
C10	<{5 Papers}, {Physical Layer, Unit Level, RFID sensor networks}>
C11	<{5 Papers}, {Physical Layer, RFID sensor networks, Wireless sensor networks (WSN)}>
C12	<{4 Papers}, { Physical Layer, Unit Level, System level, System-of-systems level, RFID sensor networks}>
C13	<{4 Papers}, {Physical Layer, RFID sensor networks, Unit level, Wireless sensor networks (WSN)}>
C14	<{4 Papers}, {Physical Layer, System level, Wireless sensor networks (WSN)}>
C15	<{4 Papers}, {Physical Layer, Wireless sensor networks (WSN), RFID}>

C16	<{3 Papers}, {Physical Layer, Unit Level, System level, System-of-systems level, Wireless sensor networks (WSN), RFID sensor networks}>
C17	<{3 Papers}, {Physical Layer, Unit Level, Wireless sensor networks (WSN), RFID}>
C18	<{3 Papers}, {Physical Layer, System Level, Wireless sensor networks (WSN), RFID }>
C19	<{2 Papers}, {Physical Layer, Unit Level, System level, System-of-systems level, RFID sensor networks, RFID, Wireless sensor networks (WSN)}>

The concepts C2, C3, C4, C5 show that the literature differs regarding the definition of the physical space, as highlighted also by the DT analysis definitions. In literature review, the digital twin has been applied at different physical levels (Tao et al., 2019) that can be summarized in: unit level (C3), system level (C4), system-of-systems level (C5). The unit level is a minimum but independent individual, which cannot be further divided such as a single piece of equipment (e.g., a product, a machine tool or robot arm). It contains a basic closed loop of data between the physical and virtual spaces with the abilities of sensing and computing. The system level can be a production system such as a production line, a shop floor, or a factory. It is characterized by self-organization, self-configuration, and self-optimization. While the system-of-systems level is characterized by enterprises' collaborations. The application of the digital twin at the system-of systems level can achieve the horizontal integration. It refers to the exchange of information across the supply chain such as resources management system, logistics, marketing, or intercompany value chains (Posada et al., 2015). For each type of level, Digital Twins can get and share data between all production factors and information systems achieving the vertical integration i.e., the integration of various Information Technology systems at different hierarchical levels. In manufacturing contexts, the literature concurs that the data type and consequently the data sources depend on the selected physical levels. Typically, a system-of-systems digital twin involve and exploit different data sources such as Internet/Users Data from CRM, E-commerce platforms (e.g., Amazon) and social networking platforms (e.g., Twitter, Facebook, LinkedIn, and YouTube), to understand user preference, and behaviours (Qi and Tao, 2018). It also involves Product data from computer-aided systems like CAD/CAM, CAE; Management data from manufacturing information systems such as MES, PDM, SCM, ERP, etc (Luo et al., 2018); Operational data from manufacturing equipment such as product data, quality data, maintenance data (Dassisti et al., 2019b); Environmental data which affects the physical equipment operations, such as environmental pressure, ambient temperature, and moisture level (Cai et al., 2017). The papers grouped in C6-C19 treat the equipping of a physical system with sensors, actuators, and embedded communication for recording real-time states (e.g., vibration, force, torque, and speed) and working conditions (e.g., environment parameters, loads, and control orders) (Ruppert et al., 2018) of the physical space. The most discussed technologies for the physical layer are RFID, RFID Sensor networks and Wireless sensor networks (WSN). RFID allows automatic identification and data capture using radio waves, a tag, and a reader (Lee and Lee, 2015). RFID Sensor networks, consisting of a very large number of nodes for monitoring and recording the physical conditions of the environment (Atzori et al., 2010). Wireless sensor networks (WSN) which consist of spatially distributed autonomous sensor-equipped devices to monitor physical or environmental conditions (Gubbi et al., 2013), (Tan and Wang, 2010). The components belonging to physical layer carry out

real-time data for the synchronization of the virtual twin with its corresponding physical twin with the capabilities of anomaly detection, prediction, prescription, and optimization.

The network layer involves connections and interactions amongst physical elements and virtual space. This layer connects all components together for sharing data and information with other connected components (Da Xu et al., 2014). The key technologies discussed in the literature review are middleware, communication protocol analysis, communication protocol/interface conversion, wireless communication, and Application Programming Interfaces (API). FCA detects 11 possible concepts, as illustrated in Figure 16 and in Table 7.

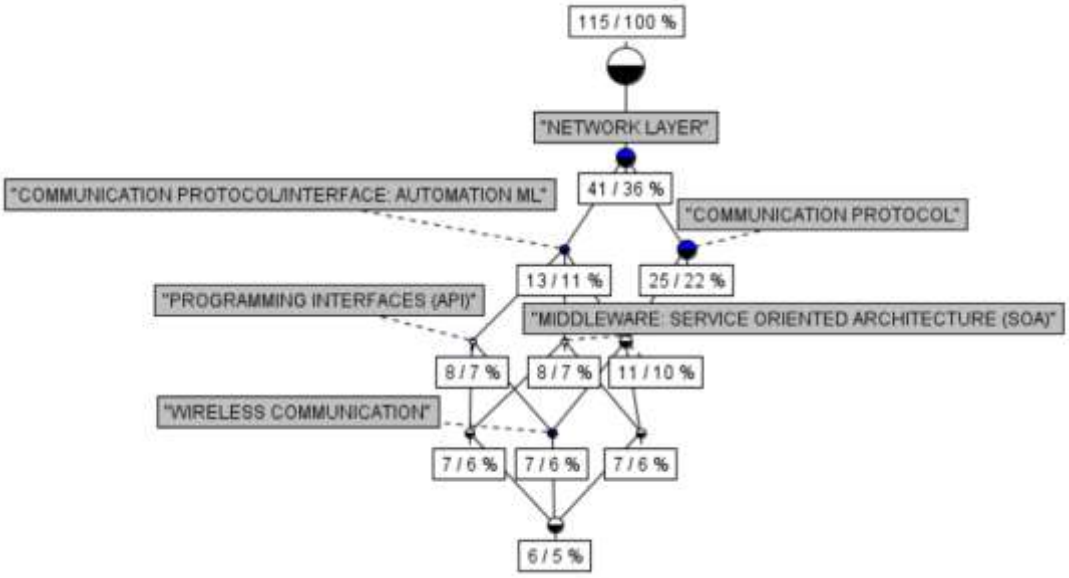


Figure 16: Lattice of Networking Layer and its Components/Technologies

Table 7: Formal Concepts of Networking Layer and its Components/Technologies

FORMAL CONCEPTS	
ID CONCEPT	<{CONCEPT EXTENTS},{CONCEPT INTENTS}>
C1	<{115 Papers}>
C2	<{41 Papers}, {Network Layer}>
C3	<{13 Papers}, {Network Layer, Communication protocol/interface: Automation ML}>
C4	<{25 Papers}, {Network Layer, Communication protocols}>
C5	<{8 Papers}, {Network Layer, Communication protocol/interface: Automation ML, Programming interface (API)}>
C6	<{8 Papers}, {Network Layer, Communication protocol/interface: Automation ML, Middleware}>
C7	<{11 Papers}, {Network Layer, Communication protocol/interface: Automation ML Communication protocol}>
C8	<{7 Papers}, {Network Layer, Communication protocol/interface: Automation ML, Programming interface (API), Middleware}>
C9	<{7 Papers}, {Network Layer, Communication protocol/interface: Automation ML, Programming interface (API), Communication protocol, Wireless Communication}>

C10	<{7 Papers}, {Network Layer, Communication protocol/interface: Automation ML, Middleware}>
C11	<{6 Papers}, {Network Layer, Communication protocol/interface: Automation ML, Middleware, Programming interface (API), Communication protocol, Wireless Communication}>

C4 concept shows, through its Extent column, the existence, in our literature review, of {25 papers} which analyse the <{Communication protocols} for the {Network Layer}>. The communication protocol allows two or more entities in the DT to transmit information to each other. OPC Unified Architecture (OPC UA) and MT-Connect protocols are the protocols more employed in digital twin applications to access to data and to transmit them in real-time (Redelinghuys et al., 2019). The studies {13 papers} in concept C3 discuss the <{Communication protocol/interface: Automation ML} in the {Network Layer}>. The communication protocol/interface conversion transforms various communication protocols/ interfaces into a unit form. AutomationML is an open standard for a data format based on XML allowing the exchange of plant engineering information (Bao et al., 2018), (Drath et al., 2008). The AutomationML is used in digital twin to model attributes related to the digital twin. The goal is to interconnect the heterogeneous toolchain of digital manufacturing (Um et al., 2017). It is used to exchange data between the Digital Twin and other systems and a methodology for communication and exchange of data (G. N. Schroeder et al., 2016), (Talkhestani et al., 2018). The studies belonging to concepts C5-C11 help to deepen the following technologies <{Middleware, Wireless communication, Application Programming Interfaces (API)} in the {Network Layer}>. The middleware is a software layer interposed between the technological and the application levels. The middleware architecture more used in the digital twin is the Service Oriented Architecture (SOA) approach. The adoption of the SOA principles allows for decomposing complex and monolithic systems into applications consisting of an ecosystem of simpler and well-defined components (Gubbi et al. 2013b). The wireless communication can connect entities in the DT wirelessly, thus improving flexibility in data transmission. The application Programming Interfaces (API) realize the communication between different software systems and models in the virtual space that represents the computing layer.

The computing layer is fundamental for computing and decisional support of digital twins. FCA detects 90 possible concepts as shown in Figure 16. A set of concepts are reported in Table 8. C2 concept demonstrates that the <{Computing Layer}> is the most addressed in our literature review {83 papers} either for the potential innovation or for the strong impact on decision support.

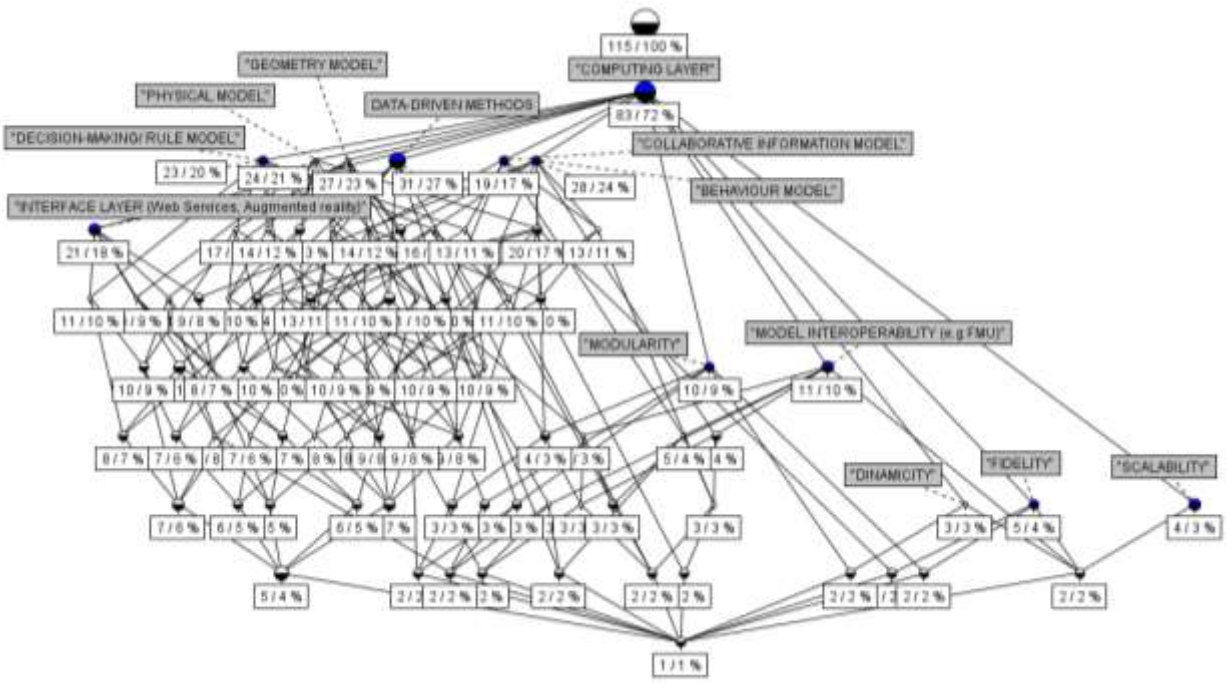


Figure 17: Lattice of Computing Layer and its Components/Technologies

Table 8: Formal Concepts of Computing Layer and its Components/Technologies

FORMAL CONCEPTS	
ID CONCEPT	<{CONCEPT EXTENTS},{CONCEPT INTENTS}>
C1	<{115 Papers}>
C2	<{83 Papers}, {Computing Layer}>
C3	<{31 Papers}, {Computing Layer, Data-driven Methods}>
C4	<{23 Papers}, {Computing Layer, Decision-making/Rule model}>
C5	<{24 Papers}, {Computing Layer, Physical model}>
C6	<{27 Papers}, {Computing Layer, Geometric model}>
C7	<{19 Papers}, {Computing Layer, Collaborative information model}>
C8	<{28 Papers}, {Computing Layer, Behaviour model}>
C9	<{10 Papers}, {Computing Layer, Decision-making/Rule model, Physical model, Geometric model, Collaborative information model, Behaviour model}>
C10	<{10 Papers}, {Computing Layer, Modularity}>
C11	<{11 Papers}, {Computing Layer, Interoperability}>
C12	<{3 Paper}, {Computing Layer, Dynamicity}>
C13	<{5 Paper}, {Computing Layer, Fidelity}>
C14	<{4 Paper}, {Computing Layer, Scalability}>

The computing layer can be perceived as a set of "layers" interconnected, which includes the following components: data (C3), models (C4-C9), and modelling features (C10-C14).

The data layer includes all different types of data previously defined in the physical layer (Uhlemann et al., 2017). This sub-layer has characteristics of heterogeneity of data and data sources, volume, and speediness. Data preparation and data analysis are the key aspects discussed in literature. The data

preparation process includes data selection, data cleaning, data modelling, data integration, and data transformation. The data analysis includes all data-driven models such as machine learning data mining, pattern evaluation, and knowledge representation involved in DT building. The studies {31 papers} grouped in C3 examine the <{Data-driven Methods} in {Computing Layer}>. Data-driven models are designed to extract knowledge from data (Y. Zhang et al., 2017), (Lee et al., 2014b). The digital twin aims to integrate data across different domains into virtual models (Kusiak, 2018). The main data-driven models used in digital twins are machine learning, neural networks, and deep learning. The machine learning refers to the ability to give computers the possibility to learn without being explicitly programmed (Clarke et al., 2009). It is classified in supervised, unsupervised (Sutharssan et al., 2015) and reinforcement learning (van Otterlo and Wiering, 2012) (Jain et al., 2017). Machine learning techniques used in digital twins are supervised and unsupervised learning (Jain et al., 2017). The supervised learning develops models based on input and output data (Tidriri et al., 2016). The supervised learning, across the digital twin, is applied for the system's failures prediction (Asimov et al., 2018), or for prediction of the remaining useful life (RUL) of the physical twin (Z. Liu et al., 2018). The unsupervised learning, instead, discovers an internal representation from input data only (Sutharssan et al., 2015), (Fahad et al., 2014). It enables discovering similar groups within data, based on clustering techniques (Xu and Wunsch, 2005), (Grira et al., 2004). In digital twins, these techniques are used for creating autonomous clusters for different working regimes to analyse machine conditions (Lee et al., 2014a), (Banerjee et al., 2017) (Ding et al., 2019). Artificial Neural Networks (ANN) and Deep Learning (DL) are computing systems that are inspired by the human brain (Zhang, 2000). The main scopes in using neural networks and deep learning in digital twin (Lee et al., 2013) are health assessment, performance prediction (Jain and Bhatnagar, 2020), fault diagnosis (Xu et al., 2019).

The digital twin requires the building and the applying digital models representing the set of resources and processes knowledge. Different tools and technologies are available for developing high-fidelity virtual models (Schleich et al., 2017). The most discussed components regard the model types (concepts C4-C9) and the modelling features (concepts C10-C14). The model types define the physics-based models and the functions of each model necessary to emulate the physical system. Physics-based models compare simulated results with known information, represented by mathematical models (Tidriri et al., 2016). A model represents a system in terms of logical and quantitative relationships that are then manipulated and changed to see how the model reacts, and thus how the system would react-if the mathematical model is a valid one (Law et al., 2000). The physics-based models are based on a set of different models to represent the structure, the behaviour, and the interactions of a physical system (Tidriri et al., 2016), (Tao et al., 2018c). The most studied models for developing a digital twin (Semeraro et al., 2019a) are summarised as follow: Geometric model {27 Papers} in concept C6; Physical model in {24 Papers}in concept C5; Behaviour model {28 Papers} in C8; Collaborative information model {19 Papers} in C7; Decision-making model {23 Papers} in C4. A geometric model reflects the geometry, the kinematics, the logic, and the interfaces of the real system (Ayani et al., 2018), (Xie et al., 2019). A geometric model defining shapes, sizes, positions and assembling machine components is presented in (Tao et al., 2018c). A physical model enables to simulate the physical properties and loads (Post et al., 2009) analysing the phenomena, such as deformation, cracking and corrosion (Tao et al., 2018c). A behaviour model describes the

way the physical system is governed by driving factors (e.g., control orders) or disturbing factors (e.g., human interferences) (Tao et al., 2018c), (Bao et al., 2018). A collaborative information model (Bao et al., 2018) defines how different components interact and simulates the collaborative behaviour among several assets. A decision-making model makes the model capable of evaluating, reasoning, and validating. It consists of variable input, algorithms and a collection of constraints and rules (Bao et al., 2018). It includes rules of constraints, associations, and deductions (Tao et al., 2018c) and it stores and analyses the running status data, then it makes decisions using machine learning algorithms.

The common features studied for modelling a digital twin concern: scalability, interoperability, fidelity, dynamicity, and modularity. According to the studies {4 Papers} grouped in C14, the scalability is the ability to provide an insight at different scales (from fine details to large systems) (Schleich et al., 2017) (Putnik et al., 2013). The studies {11 Papers} in C11 define the interoperability as the ability to convert, to combine, and to establish equivalence between different model representations (H. Zhang et al., 2017). The model interoperability is a critical aspect for the exchange of dynamic models and for Co-Simulation. Functional Mock-Up Interface (FMU) standard is commonly used in digital twins to solve this problem (Negri et al., 2019) (Schluse et al., 2018). FMU is an open standard for exchanging dynamical simulation models between different tools in a standardized format and for co-simulation (Blochwitz et al., 2011). FMI standard specifies two different kinds of FMUs: (1) Model Exchange (ME) – ME FMUs; (2) Co-Simulation (CS) – CS FMUs. The model fidelity (concept C13) describes the closeness to the physical product (Schleich et al., 2017) while the model dynamicity (concept C12) is the ability to reflect real time the physical process and modify autonomously itself if the physical system changes. This crucial issue concerns the convergence of the physical world with its digital counterpart (Weyer et al., 2016). According to the {10 Papers} clustered in C10, the modularity is the ability to integrate, to add, or to replace models (Guo et al., 2018). Two modular approaches have been developed in (Guo et al., 2018), (Semeraro et al., 2019a). The idea behind this approach is to use and especially re-use predefined functional units (Semeraro et al., 2019a), that are systematically developed and logically interlinked for the configuration of a holistic manufacturing system (Stark et al., 2017), (Negri et al., 2019). Virtual (VR) or augmented (AR) reality technologies can be integrated in digital twins to create interactive and immersive environments (G. Schroeder et al., 2016) enabling direct interactions between the digital twin and final users.

As a summary of the analysis discussed above, the digital twin paradigm is summarised in Figure 18 to depict the contexts, the phases of the life cycle (design, production, and service), the functions of the Digital Twin for each life cycle phase, the architecture layers, and the main components of each architecture layer.

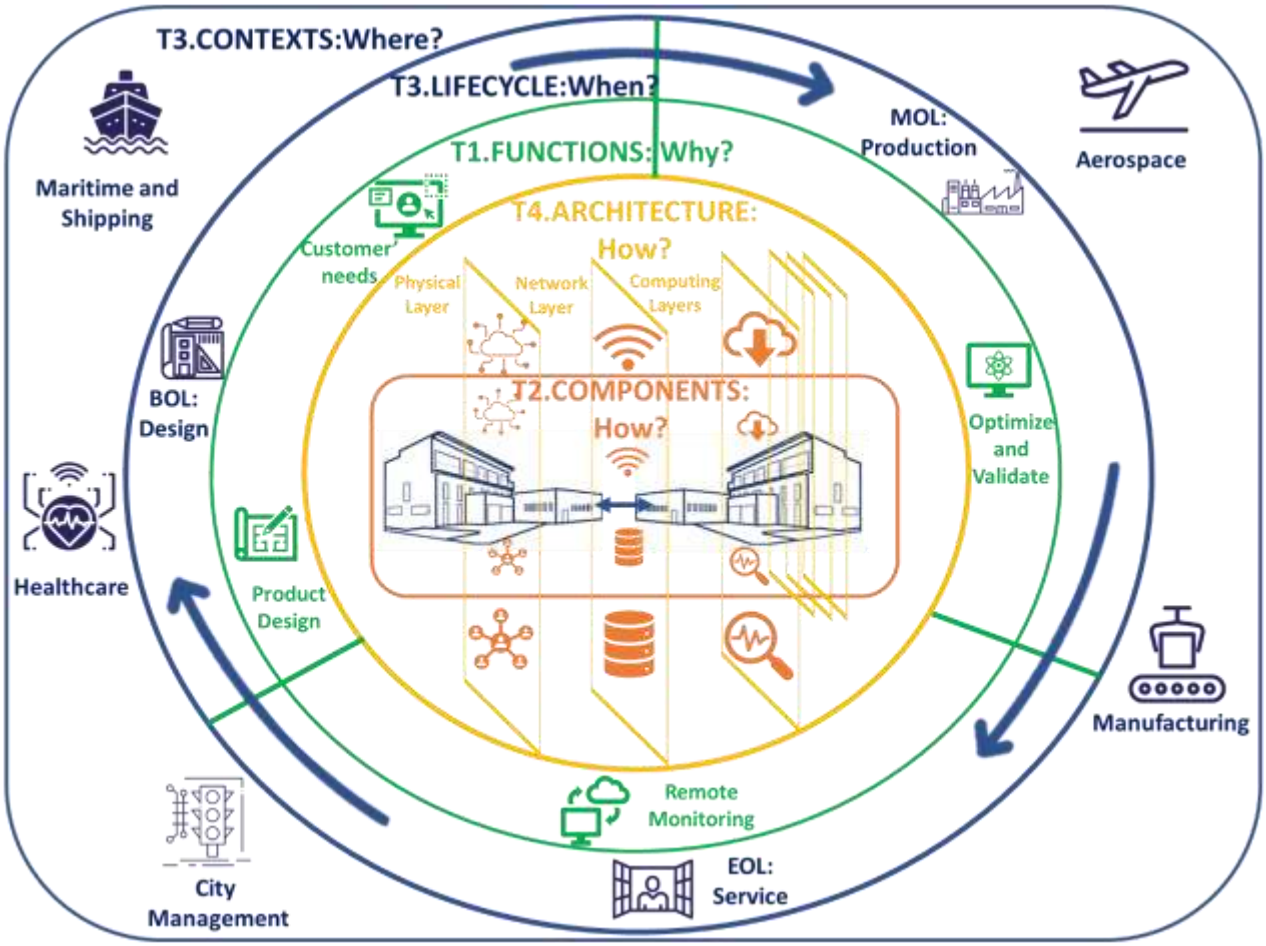


Figure 18: The Digital Twin Paradigm (Semeraro, 2020)

4 Conclusion and Research Challenges

The present study forms a literature review that led to a digital twin paradigm aiming at assessing which are the application contexts, the life cycle phases, the functionalities, the architectures, and the components of existing digital twins. The paper aims at providing a detailed picture of the main features of existing scientific research on DT's, stressing on the different application domains and the related technologies. The idea of Digital Twin as a "virtual" image of the reality constantly synchronized with the real operating scenario is accurately presented and described in section 1 in all its physical and logical aspects. This literature review tries to answer different research questions at different level namely: DT definition, application contexts, life cycle phases, functions, architectures, and components. Section 2 investigates on different DT definition provided in literature to address the research question: 'What is a Digital Twin?'. The application contexts in section 3.1 and life cycle phases in section 3.2 focus on establishing the baseline of the Digital Twin paradigm by trying to reply to the research questions: 'Where is appropriate to use a Digital Twin?', 'Who is doing Digital Twins?', 'When has a Digital Twin to be developed?'. A digital twin may enable companies and organisation to predict outcomes, design, and build better products, and better serve their customers (Madni et al., 2019). To that point, the third research question ('Why should a Digital Twin be used?'), in section 3.2, analyses the main functions of a digital twin for each life cycle phase. The fourth

research question investigated in section 3.3 ('How to design and implement a Digital Twin?') tries to define the DT architecture and analyse the employed components/technologies for implementing digital twins. The review results are summarised in the summary table below.

Table 9: Summary table of the review results (5W1H)

Research Question	Results
'What is a Digital Twin?' **DIGITAL TWIN DEFINITION**	*"A set of **adaptive models** that emulate the **behaviour** of a physical system in a **virtual system** getting **real time data** to update itself along its life cycle. The digital twin replicates the physical system to predict failures and opportunities for changing, to prescribe real time actions for **optimizing and/or mitigating** unexpected events observing and evaluating the operating profile system".*
'Where is appropriate to use a Digital Twin?' **DIGITAL TWIN CONTEXTS and USE CASES**	1. Healthcare • Improving operational efficiency of healthcare operations 2. Maritime and Shipping • Design customization 3. Manufacturing • Product development and predictive manufacturing 4. City Management • Modelling and simulation of smart cities 5. Aerospace • Predictive analytics to foresee future aircraft problems
'Who is doing Digital Twins?' **DIGITAL TWIN PLATFORMS**	GE Predix; SIEMENS PLM; Microsoft Azure; IBM Watson; PTC Thing Worx; Aveva; Twin Thread; DNV-GL; Dassault 3D Experience; Sight Machine; Oracle Cloud.
'When has a Digital Twin to be developed?' **DIGITAL TWIN LIFE CYCLE** **'Why should a Digital Twin be used?'** **DIGITAL TWIN FUNCTIONS**	1. In design phase • The digital twin is used to help designers to configure and validate more quickly the product development accurately interpreting the market demands and the customer preferences. 2. In production phase • The digital twin shows a great potential in real-time process control and optimization, as well as accurate prediction. 3. In service phase • The digital twin can monitor the health of a product, perform diagnosis and prognosis.
'How to design and implement a Digital Twin?' **DIGITAL TWIN ARCHITECTURE AND COMPONENTS**	1. The Physical layer involves various subsystems and sensory devices that collect data and working parameters. 2. The Network layer connects the physical to the virtual, sharing of data and information. 3. The Computing layer consists of virtual models emulating the corresponding physical entities.

The analysis of the digital twin definitions, features, tools, and methodologies was done based on the text mining techniques and Formal Concept Analysis (FCA). The application contexts, life cycle phases, functionalities, architectures, and components are discussed and organised in a unique paradigm summarised in Figure 18. Balanced against the many advantages that the digital twin can bring, there are several challenges to be overcome. For the grace of the results summarised in table 9, we try to define what are the main research challenges of implementing a Digital Twin.

Research Challenge #1.
DT APPLICATION CONTEXTS

According to section 3.1, the existing digital twin applications refer to specific and traditional contexts: *Healthcare; Maritime and Shipping; Manufacturing; City Management; Aerospace*. Humans play an important role in the Digital Twin applications especially in manufacturing contexts. Human interaction is the one key challenge in the development and implementation of DT in the manufacturing application. While some low-level operations can be autonomously achieved without human intervention, many decision-making activities are still sustained by many manual operations based upon human interactions. The research challenge is to form and design cognitive digital twin able to interoperate with other digital twins and humans in a seamless way whatever happens during their interactions. Furthermore, the digital twins could virtually help look beyond the current industrial model of extracting, producing, consuming, and disposing. The digital twin can support companies to move from a linear system to a 'circular economy' system that considers almost zero waste production and pollution, keeps products and materials in the recycling loop longer and helps regenerate natural systems. While scientific literature has analysed the adoption of DT in the optimization of products life cycle, few contributions have yet focused on the exploitation of DT to assess and improve the sustainability performances of whole value chains (Barni et al. 2018). Significant research efforts need to be made on the application of a digital twin for improving the sustainability performances in each application context.

Research Challenge #2.
DT LIFE CYCLE AND FUNCTIONS

The DT presents an exciting possibility of real-time simulation for product life cycle that can be divided into three phases: design, production, and service. The digital twin can be developed in each life cycle phase fulfilling different functions. The digital twin can potentially help to integrate even the entire supply chain, throughout all phases of product life cycle. FCA results in section 3.2 demonstrates that most digital twin applications refer to a single phase of a product life cycle. There are currently relatively few applications of digital twin for supporting the entire supply chain and the network enterprises. Digital Twin applications are mainly developed for prediction purposes and used for decision-making support. The digital twin could connect products, persons, machines, and enterprises within the virtual space. However, this aspect holds its own challenges at the present state. The ability to collect, aggregate and exchange data and information between different suppliers, manufacturers and customers could present interoperability issues. As the DT can potentially integrate data from the lifetime of a product seamlessly, several key research challenges concern the definition of standards and communication protocols to ensure interoperability of multiple digital

twins with each other (Platenius-Mohr et al. 2020). The development of standards-based interoperability is important and challenging at the same time for the evolution of digital twin applications along the entire life cycle.

Research Challenge #3.
DT ARCHITECTURE AND COMPONENTS/TECHNOLOGIES

The DT consists of a set of models with complex structures and behaviour, that reflect the real-time operations of the physical system. Modelling a digital copy of a physical system to perform real-time validation and optimization is quite complex and thus needs a big amount of data. A digital twin can be a model of a component, a system of components or a system of systems. The digital twin requires the construction and application of accurate models of reality. Modelling the reality in a digital twin is a complex task, as it involves sensors, multifunctional models, multisource data, services, etc. At present, it is difficult to build an accurate model for a DT using traditional approaches due to the complexity of real systems. The lack of an univocal reference architecture leads at developing Digital Twin solutions using different technologies, interfaces, and communication protocols, models and data as assessed in section 3.3. Standard Digital Twin solutions should be developed to provide design criteria and design constraints where reference architectural aspects, reference information model and communication protocols are clearly defined (Lu et al. 2020). The modularization design principle needs to be explored to improve the modelling efficiency. This would enable to improve the flexibility and reusability of standard DT-solutions towards different applications. Modular approaches need to be explored for the construction of flexible DT solutions, facilitating new DT applications.

Next steps of the present research will be related to the definition of a new approach for building a digital twin by exploring the modularity feature, which is still one of the most challenging research issues. The idea that we will explore is the definition of a modelling approach that allows to derive a criterion to self-detect modelling constructs that can be used (and re-used) to create digital models of different systems or processes.

Authors' contributions:

Author CS was responsible for the study conduction and assimilating the literature to select the final sample, also defining the methodological approach; Authors ML, HP, MD assessed the quality of the included studies; CS synthesised the literature and wrote a first draft of the manuscript; ML, HP, MD contributed to the final version and provided several suggestions to improve the quality of the systematic literature review and the final research challenges. All authors have read and agreed to the paper being submitted in the present form.

Declaration of Competing Interest

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

Acknowledgements

This work was supported by the Italian Ministry of Education, University and Research under the Programme "Department of Excellence" Legge 232/2016 (Grant No. CUP - D94I18000260001)

References

Abramovici, M., Göbel, J.C., Savarino, P., 2017. Reconfiguration of smart products during their use phase based on virtual product twins. CIRP Annals 66, 165–168.

Agrawal, R., Imieliński, T., Swami, A., 1993. Mining association rules between sets of items in large databases, in: Acm Sigmod Record. ACM, pp. 207–216.

Alam, K.M., El Saddik, A., 2017. C2PS: A digital twin architecture reference model for the cloud-based cyber-physical systems. IEEE Access 5, 2050–2062.

Alcácer, V., Cruz-Machado, V., 2019. Scanning the Industry 4.0: A Literature Review on Technologies for Manufacturing Systems. Engineering Science and Technology, an International Journal.

Arrichiello, V., Gualeni, P., 2019. Systems engineering and digital twin: a vision for the future of cruise ships design, production and operations. International Journal on Interactive Design and Manufacturing (IJIDeM) 1–8.

Asimov, R.M., Chernoshey, S.V., Kruse, I., Osipovich, V.S., 2018. Digital twin in the Analysis of a Big Data.

Atzori, L., Iera, A., Morabito, G., 2010. The internet of things: A survey. Computer networks 54, 2787–2805.

Autiosalo, J., 2018. Platform for industrial internet and digital twin focused education, research, and innovation: Ilmatar the overhead crane, in: 2018 IEEE 4th World Forum on Internet of Things (WF-IoT). Presented at the 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), pp. 241–244. https://doi.org/10.1109/WF-IoT.2018.8355217

Ayani, M., Ganebäck, M., Ng, A.H.C., 2018. Digital Twin: Applying emulation for machine reconditioning. Procedia CIRP, 51st CIRP Conference on Manufacturing Systems 72, 243–248. https://doi.org/10.1016/j.procir.2018.03.139

Bachelor, G., Brusa, E., Ferretto, D., Mitschke, A., 2019. Model-Based Design of Complex Aeronautical Systems Through Digital Twin and Thread Concepts. IEEE Systems Journal.

Banerjee, A., Dalal, R., Mittal, S., Joshi, K.P., 2017. Generating digital twin models using knowledge graphs for industrial production lines. UMBC Information Systems Department.

Bao, J., Guo, D., Li, J., Zhang, J., 2018. The modelling and operations for the digital twin in the context of manufacturing. Enterprise Information Systems 0, 1–23. https://doi.org/10.1080/17517575.2018.1526324

Barni, Andrea, Alessandro Fontana, Silvia Menato, Marzio Sorlini, and Luca Canetta. 2018. 'Exploiting the Digital Twin in the Assessment and Optimization of Sustainability Performances'. In 2018 International Conference on Intelligent Systems (IS), 706–13. IEEE.

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

Blochwitz, T., Otter, M., Arnold, M., Bausch, C., Elmqvist, H., Junghanns, A., Mauß, J., Monteiro, M., Neidhold, T., Neumerkel, D., 2011. The functional mockup interface for tool independent exchange of simulation models, in: Proceedings of the 8th International Modelica Conference; March 20th-22nd; Technical Univeristy; Dresden; Germany. Linköping University Electronic Press, pp. 105–114.

Blume, M., Koch, N., Imtiaz, J., Flatt, H., Jasperneite, J., Schleipen, M., Sauer, O., Dosch, S., 2014. An OPC-UA based approach for dynamic-configuration of security credentials and integrating a vendor independent digital product memory, in: Jahreskolloquium Kommunikation in Der Automation (KommA 2014).

Bohlin, R., Hagmar, J., Bengtsson, K., Lindkvist, L., Carlson, J.S., Söderberg, R., 2017. Data flow and communication framework supporting digital twin for geometry assurance. Presented at the ASME International Mechanical Engineering Congress and Exposition, Proceedings (IMECE). https://doi.org/10.1115/IMECE2017-71405

Boje, C., Guerriero, A., Kubicki, S., Rezgui, Y., 2020. Towards a semantic Construction Digital Twin: Directions for future research. Automation in Construction 114, 103179.

Borangiu, T., Trentesaux, D., Thomas, A., Leitão, P., Barata, J., 2019. Digital transformation of manufacturing through cloud services and resource virtualization. Computers in Industry 108, 150–162. https://doi.org/10.1016/j.compind.2019.01.006

Boschert, S., Heinrich, C., Rosen, R., 2018. Next generation digital twin, in: Proc. TMCE. Las Palmas de Gran Canaria, Spain, pp. 209–217.

Boschert, S., Rosen, R., 2016. Digital twin—the simulation aspect, in: Mechatronic Futures. Springer, pp. 59–74.

Brenner, B., Hummel, V., 2017. Digital Twin as Enabler for an Innovative Digital Shopfloor Management System in the ESB Logistics Learning Factory at Reutlingen - University. Procedia Manufacturing, 7th Conference on Learning Factories, CLF 2017 9, 198–205. https://doi.org/10.1016/j.promfg.2017.04.039

Brettel, M., Friederichsen, N., Keller, M., Rosenberg, M., 2014. How virtualization, decentralization and network building change the manufacturing landscape: An Industry 4.0 Perspective. International journal of mechanical, industrial science and engineering 8, 37–44.

Bruynseels, K., Santoni de Sio, F., van den Hoven, J., 2018. Digital twins in health care: ethical implications of an emerging engineering paradigm. Frontiers in genetics 9, 31.

Büchi, G., Cugno, M., Castagnoli, R., 2020. Smart factory performance and Industry 4.0. Technological Forecasting and Social Change 150, 119790. https://doi.org/10.1016/j.techfore.2019.119790

Cai, Y., Starly, B., Cohen, P., Lee, Y.-S., 2017. Sensor data and information fusion to construct digital-twins virtual machine tools for cyber-physical manufacturing. Procedia Manufacturing 10, 1031–1042.

Cheng, J., Zhang, H., Tao, F., Juang, C.-F., 2020. DT-II: Digital twin enhanced Industrial Internet reference framework towards smart manufacturing. Robotics and Computer-Integrated Manufacturing 62, 101881.
Ciavotta, M., Alge, M., Menato, S., Rovere, D., Pedrazzoli, P., 2017. A microservice-based middleware for the digital factory. Procedia Manufacturing 11, 931–938.
Cimino, C., Negri, E., Fumagalli, L., 2019. Review of digital twin applications in manufacturing. Computers in Industry 113, 103130.
Clarke, B., Fokoue, E., Zhang, H.H., 2009. Principles and theory for data mining and machine learning. Springer Science & Business Media.
Da Xu, L., He, W., Li, S., 2014. Internet of things in industries: A survey. IEEE Transactions on industrial informatics 10, 2233–2243.
Damiani, L., Demartini, M., Giribone, P., Maggiani, M., Revetria, R., Tonelli, F., 2018. Simulation and Digital Twin Based Design of a Production Line: A Case Study.", in: Proceedings of the International MultiConference of Engineers and Computer Scientists.
Dassisti, M., De Nicolò, M., 2012. Enterprise integration and economical crisis for mass craftsmanship: a case study of an Italian furniture company, in: OTM Confederated International Conferences" On the Move to Meaningful Internet Systems". Springer, pp. 113–123.
Dassisti, M., Giovannini, A., Merla, P., Chimienti, M., Panetto, H., 2019a. An approach to support Industry 4.0 adoption in SMEs using a core-metamodel. Annual Reviews in Control 47, 266–274.
Dassisti, M., Giovannini, A., Merla, P., Chimienti, M., Panetto, H., 2017. Hybrid Production-System Control-Architecture for Smart Manufacturing, in: OTM Confederated International Conferences" On the Move to Meaningful Internet Systems". Springer, pp. 5–15.
Dassisti, M., Panetto, H., Lezoche, M., Merla, P., Semeraro, C., Giovannini, A., Chimienti, M., 2017b. Industry 4.0 paradigm: The viewpoint of the small and medium enterprises, in: 7th International Conference on Information Society and Technology, ICIST 2017. pp. 50–54.
Dassisti, M., Semeraro, C., 2018. Smart Sustainable Manufacturing: a new Laboratory-Factory concept to test Industry 4.0 principles, in: 8th International Conference on Information Society and Technology, ICIST 2018.
Dassisti, M., Semeraro, C., Chimenti, M., 2019b. Hybrid Exergetic Analysis-LCA approach and the Industry 4.0 paradigm: Assessing Manufacturing Sustainability in an Italian SME. Procedia Manufacturing 33, 655–662.
Ding, K., Chan, F.T.S., Zhang, X., Zhou, G., Zhang, F., 2019. Defining a Digital Twin-based Cyber-Physical Production System for autonomous manufacturing in smart shop floors. International Journal of Production Research 0, 1–20. https://doi.org/10.1080/00207543.2019.1566661
Drath, R., Horch, A., 2014. Industrie 4.0: Hit or hype?[industry forum]. IEEE industrial electronics magazine 8, 56–58.
Drath, R., Luder, A., Peschke, J., Hundt, L., 2008. AutomationML-the glue for seamless automation engineering, in: 2008 IEEE International Conference on Emerging Technologies and Factory Automation. IEEE, pp. 616–623.
Estefan, J.A., 2007. Survey of model-based systems engineering (MBSE) methodologies. Incose MBSE Focus Group 25, 1–12.
Fahad, A., Alshatri, N., Tari, Z., Alamri, A., Khalil, I., Zomaya, A.Y., Foufou, S., Bouras, A., 2014. A survey of clustering algorithms for big data: Taxonomy and empirical analysis. IEEE transactions on emerging topics in computing 2, 267–279.
Fischer, K., Heintel, M., 2017. Prüfung einer Konsistenz zwischen Referenzdaten eines Fertigungsobjektes und Daten eines digitalen Zwillings des Fertigungsobjektes. DE102015217855A1.
Garetti, M., Rosa, P., Terzi, S., 2012. Life cycle simulation for the design of product–service systems. Computers in Industry 63, 361–369.
Glaessgen, E., Stargel, D., 2012. The digital twin paradigm for future NASA and US Air Force vehicles, in: 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA. p. 1818.
Graessler, I., Poehler, A., 2017. Integration of a digital twin as human representation in a scheduling procedure of a cyber-physical production system, in: Industrial Engineering and Engineering Management (IEEM), 2017 IEEE International Conference On. IEEE, pp. 289–293.
Greif, T., Stein, N., Flath, C.M., 2020. Peeking into the void: Digital twins for construction site logistics. Computers in Industry 121, 103264. https://doi.org/10.1016/j.compind.2020.103264
Grieves, M., 2014. Digital twin: manufacturing excellence through virtual factory replication. White paper.
Grieves, M., Vickers, J., 2017. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems, in: Transdisciplinary Perspectives on Complex Systems. Springer, pp. 85–113.
Grira, N., Crucianu, M., Boujemaa, N., 2004. Unsupervised and semi-supervised clustering: a brief survey. A review of machine learning techniques for processing multimedia content 1, 9–16.
Gubbi, J., Buyya, R., Marusic, S., Palaniswami, M., 2013. Internet of Things (IoT): A vision, architectural elements, and future directions. Future generation computer systems 29, 1645–1660.
Guo, J., Zhao, N., Sun, L., Zhang, S., 2018. Modular based flexible digital twin for factory design. Journal of Ambient Intelligence and Humanized Computing 1–12.
Haag, S., Anderl, R., 2018. Digital twin – Proof of concept. Manufacturing Letters, Industry 4.0 and Smart Manufacturing 15, 64–66. https://doi.org/10.1016/j.mfglet.2018.02.006
Hearst, M., 2003. What is text mining. SIMS, UC Berkeley 5.
Hermann, M., Pentek, T., Otto, B., 2016. Design principles for industrie 4.0 scenarios, in: System Sciences (HICSS), 2016 49th Hawaii International Conference On. IEEE, pp. 3928–3937.
Hershey, J.E., Wheeler, F.W., Nielsen, M.C., Johnson, C.D., Dell'Anno, M.J., JOYKUTTI, J., 2017. Digital twin of twinned physical system. US20170286572A1.

International Council on Systems Engineering (INCOSE), 2007. International Council on Systems Engineering (INCOSE), Systems Engineering Vision 2020, Version 2.03, TP-2004-004-02, September 2007. [WWW Document]. URL http://www.omg.org/spec/ OCL/

Jain, R., Bhatnagar, R., 2020. Applications of Machine Learning in Cyber Security - A Review and a Conceptual Framework for a University Setup. Advances in Intelligent Systems and Computing 921, 599–608. https://doi.org/10.1007/978-3-030-14118-9_60

Jain, S., Shao, G., Shin, S.-J., 2017. Manufacturing data analytics using a virtual factory representation. International Journal of Production Research 55, 5450–5464.

Jones, D., Snider, C., Nassehi, A., Yon, J., Hicks, B., 2020. Characterising the Digital Twin: A systematic literature review. CIRP Journal of Manufacturing Science and Technology.

Kagermann, H., Helbig, J., Hellinger, A., Wahlster, W., 2013. Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Securing the future of German manufacturing industry; final report of the Industrie 4.0 Working Group. Forschungsunion.

Kaur, M.J., Mishra, V.P., Maheshwari, P., 2020. The Convergence of Digital Twin, IoT, and Machine Learning: Transforming Data into Action, in: Farsi, M., Daneshkhah, A., Hosseinian-Far, A., Jahankhani, H. (Eds.), Digital Twin Technologies and Smart Cities, Internet of Things. Springer International Publishing, Cham, pp. 3–17. https://doi.org/10.1007/978-3-030-18732-3_1

Khajavi, Siavash H., Naser Hossein Motlagh, Alireza Jaribion, Liss C. Werner, and Jan Holmström. 2019. 'Digital Twin: Vision, Benefits, Boundaries, and Creation for Buildings'. IEEE Access 7: 147406–19.

Kritzinger, W., Karner, M., Traar, G., Henjes, J., Sihn, W., 2018. Digital Twin in manufacturing: A categorical literature review and classification. IFAC-PapersOnLine 51, 1016–1022.

Kusiak, A., 2018. Smart manufacturing. International Journal of Production Research 56, 508–517.

Law, A.M., Kelton, W.D., Kelton, W.D., 2000. Simulation modeling and analysis. McGraw-Hill New York.

Lee, C.G., Park, S.C., 2014. Survey on the virtual commissioning of manufacturing systems. Journal of Computational Design and Engineering 1, 213–222.

Lee, H., Kim, T., 2018. Smart factory use case model based on digital twin. ICIC Express Letters, Part B: Applications 9, 931–936. https://doi.org/10.24507/icicelb.09.09.931

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

Lee, J., Bagheri, B., Kao, H.-A., 2015. A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manufacturing Letters 3, 18–23.

Lee, J., Kao, H.-A., Yang, S., 2014a. Service innovation and smart analytics for industry 4.0 and big data environment. Procedia Cirp 16, 3–8.

Lee, J., Lapira, E., Yang, S., Kao, A., 2013. Predictive manufacturing system-Trends of next-generation production systems. IFAC Proceedings Volumes 46, 150–156.

Lee, J., Wu, F., Zhao, W., Ghaffari, M., Liao, L., Siegel, D., 2014b. Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications. Mechanical systems and signal processing 42, 314–334.

Leng, J., Zhang, H., Yan, D., Liu, Q., Chen, X., Zhang, D., 2019. Digital twin-driven manufacturing cyber-physical system for parallel controlling of smart workshop. Journal of Ambient Intelligence and Humanized Computing 10, 1155–1166.

Leng, J., Zhang, H., Yan, D., Liu, Q., Chen, X., Zhang, D., 2018. Digital twin-driven manufacturing cyber-physical system for parallel controlling of smart workshop. Journal of Ambient Intelligence and Humanized Computing 1–12.

Lezoche, M., Panetto, H., 2018. Cyber-Physical Systems, a new formal paradigm to model redundancy and resiliency. Enterprise Information Systems 1–22.

Liau, Y., Lee, H., Ryu, K., 2018. Digital Twin concept for smart injection molding. Presented at the IOP Conference Series: Materials Science and Engineering. https://doi.org/10.1088/1757-899X/324/1/012077

Liu, Q., Zhang, H., Leng, J., Chen, X., 2018. Digital twin-driven rapid individualised designing of automated flow-shop manufacturing system. International Journal of Production Research 1–17.

Liu, Z., Meyendorf, N., Mrad, N., 2018. The role of data fusion in predictive maintenance using digital twin. Presented at the AIP Conference Proceedings. https://doi.org/10.1063/1.5031520

Ljubljana, B.L., University of, 2005. Data Mining [WWW Document]. URL https://orange.biolab.si/ (accessed 10.12.20).

Luo, W., Hu, T., Zhu, W., Tao, F., 2018. Digital twin modeling method for CNC machine tool, in: 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC). Presented at the 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), pp. 1–4. https://doi.org/10.1109/ICNSC.2018.8361285

Ma, Y., Borrelli, F., Hencey, B., Coffey, B., Bengea, S., Haves, P., 2011. Model predictive control for the operation of building cooling systems. IEEE Transactions on control systems technology 20, 796–803.

Macchi, M., Roda, I., Negri, E., Fumagalli, L., 2018. Exploring the role of Digital Twin for Asset Life cycle Management. IFAC-PapersOnLine 51, 790–795.

Madni, A.M., Madni, C.C., Lucero, S.D., 2019. Leveraging digital twin technology in model-based systems engineering. Systems 7, 7.

Mandolla, C., Petruzzelli, A.M., Percoco, G., Urbinati, A., 2019. Building a digital twin for additive manufacturing through the exploitation of blockchain: A case analysis of the aircraft industry. Computers in Industry 109, 134–152. https://doi.org/10.1016/j.compind.2019.04.011

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., Dean, J., 2013. Distributed representations of words and phrases and their compositionality. arXiv preprint arXiv:1310.4546.

Mohammadi, N., Taylor, J.E., 2017. Smart city digital twins, in: 2017 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE, pp. 1–5.
Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., Sauer, O., Schuh, G., Sihn, W., Ueda, K., 2016. Cyber-physical systems in manufacturing. Cirp Annals 65, 621–641.
Negri, E., Fumagalli, L., Cimino, C., Macchi, M., 2019. FMU-supported simulation for CPS Digital Twin. Procedia Manufacturing, 7th International conference on Changeable, Agile, Reconfigurable and Virtual Production (CARV2018) 28, 201–206. https://doi.org/10.1016/j.promfg.2018.12.033
Negri, E., Fumagalli, L., Macchi, M., 2017. A Review of the Roles of Digital Twin in CPS-based Production Systems. Procedia Manufacturing, 27th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017, 27-30 June 2017, Modena, Italy 11, 939–948. https://doi.org/10.1016/j.promfg.2017.07.198
Nikolakis, N., Alexopoulos, K., Xanthakis, E., Chryssolouris, G., 2018. The digital twin implementation for linking the virtual representation of human-based production tasks to their physical counterpart in the factory-floor. International Journal of Computer Integrated Manufacturing 1–12.
Osterrieder, P., Budde, L., Friedli, T., 2019. The smart factory as a key construct of industry 4.0: A systematic literature review. International Journal of Production Economics. https://doi.org/10.1016/j.ijpe.2019.08.011
Onosato, M., Iwata, K., 1993. Development of a virtual manufacturing system by integrating product models and factory models. CIRP annals 42, 475–478.
Panetto, H., Iung, B., Ivanov, D., Weichhart, G., Wang, X., 2019. Challenges for the cyber-physical manufacturing enterprises of the future. Annual Reviews in Control 47, 200–213.
Park, K.T., Nam, Y.W., Lee, H.S., Im, S.J., Noh, S.D., Son, J.Y., Kim, H., 2019. Design and implementation of a digital twin application for a connected micro smart factory. International Journal of Computer Integrated Manufacturing 32, 596–614. https://doi.org/10.1080/0951192X.2019.1599439
Pirola, F., Boucher, X., Wiesner, S., Pezzotta, G., 2020. Digital technologies in product-service systems: a literature review and a research agenda. Computers in Industry 123, 103301. https://doi.org/10.1016/j.compind.2020.103301
Platenius-Mohr, Marie, Somayeh Malakuti, Sten Grüner, Johannes Schmitt, and Thomas Goldschmidt. 2020. 'File-and API-Based Interoperability of Digital Twins by Model Transformation: An IIoT Case Study Using Asset Administration Shell'. Future Generation Computer Systems 113: 94–105.
Ponomarev, K., Kudryashov, N., Popelnukha, N., Potekhin, V., 2017. Main principals and issues of digital twin development for complex technological processes. Presented at the Annals of DAAAM and Proceedings of the International DAAAM Symposium, pp. 523–528. https://doi.org/10.2507/28th.daaam.proceedings.074
Posada, J., Toro, C., Barandiaran, I., Oyarzun, D., Stricker, D., de Amicis, R., Pinto, E.B., Eisert, P., Döllner, J., Vallarino, I., 2015. Visual computing as a key enabling technology for industrie 4.0 and industrial internet. IEEE computer graphics and applications 35, 26–40.
Post, J., Groen, M., Klaseboer, G., 2009. PHYSICAL MODEL BASED DIGITAL TWINS IN MANUFACTURING PROCESSES. Optics Letters 34, 1982–4.
Prasad, B., 1989. Robotics and Factories of the Future. Springer.
Putnik, G., Sluga, A., ElMaraghy, H., Teti, R., Koren, Y., Tolio, T., Hon, B., 2013. Scalability in manufacturing systems design and operation: State-of-the-art and future developments roadmap. CIRP Annals 62, 751–774.
Qi, Q., Tao, F., 2018. Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison. IEEE Access 6, 3585–3593.
Qi, Q., Tao, F., Zuo, Y., Zhao, D., 2018. Digital Twin Service towards Smart Manufacturing. Procedia CIRP, 51st CIRP Conference on Manufacturing Systems 72, 237–242. https://doi.org/10.1016/j.procir.2018.03.103
Redelinghuys, A., Basson, A., Kruger, K., 2019. A six-layer digital twin architecture for a manufacturing cell. Studies in Computational Intelligence 803, 412–423. https://doi.org/10.1007/978-3-030-03003-2_32
Ríos, J., Hernández, J.C., Oliva, M., Mas, F., 2015. Product Avatar as Digital Counterpart of a Physical Individual Product: Literature Review and Implications in an Aircraft., in: ISPE CE. pp. 657–666.
Rodič, B., 2017. Industry 4.0 and the New Simulation Modelling Paradigm. Organizacija 50, 193–207. https://doi.org/10.1515/orga-2017-0017
Rosen, R., von Wichert, G., Lo, G., Bettenhausen, K.D., 2015. About the importance of autonomy and digital twins for the future of manufacturing. IFAC-PapersOnLine 48, 567–572.
Ruppert, T., Jaskó, S., Holczinger, T., Abonyi, J., 2018. Enabling Technologies for Operator 4.0: A Survey. Applied Sciences 8, 1650.
Schleich, B., ANWER, N., Mathieu, L., Wartzack, S., 2017. Shaping the digital twin for design and production engineering. CIRP Annals - Manufacturing Technology 66. https://doi.org/10.1016/j.cirp.2017.04.040
Schluse, M., Atorf, L., Rossmann, J., 2017. Experimentable digital twins for model-based systems engineering and simulation-based development, in: Systems Conference (SysCon), 2017 Annual IEEE International. IEEE, pp. 1–8.
Schluse, M., Priggemeyer, M., Atorf, L., Rossmann, J., 2018. Experimentable Digital Twins—Streamlining Simulation-Based Systems Engineering for Industry 4.0. IEEE Transactions on Industrial Informatics 14, 1722–1731.
Schroeder, G., Steinmetz, C., Pereira, C.E., Muller, I., Garcia, N., Espindola, D., Rodrigues, R., 2016. Visualising the digital twin using web services and augmented reality, in: 2016 IEEE 14th International Conference on Industrial Informatics (INDIN). IEEE, pp. 522–527.
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, 12–17.

Semeraro, C., Lezoche, M., Panetto, H., Dassisti, M., Cafagna, S., 2019a. Data-driven pattern-based constructs definition for the digital transformation modelling of collaborative networked manufacturing enterprises, in: Working Conference on Virtual Enterprises. Springer, pp. 507–515.

Semeraro, C., Panetto, H., Lezoche, M., Dassisti, M., Cafagna, S., 2019b. A Monitoring Strategy for Industry 4.0: Master Italy srl Case Study. INSIGHT-International Council on Systems Engineering (INCOSE) 22, 20–22.

Semeraro, C. 2020. Contribution to the Formalization of Data-Driven Invariant Modelling Constructs of Cyber-Physical Systems, PhD thesis.

Škopljanac-Mačina, F., Blašković, B., 2014. Formal concept analysis–overview and applications. Procedia Engineering 69, 1258–1267.

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, 137–140.

Song, Z., Canedo, A.M., 2016. Digital twins for energy efficient asset maintenance. US20160247129A1.

Stark, R., Kind, S., Neumeyer, S., 2017. Innovations in digital modelling for next generation manufacturing system design. CIRP Annals 66, 169–172.

Suh, N.P., 1984. The future of the factory. Robotics and Computer-Integrated Manufacturing 1, 39–49.

Sun, H., Li, C., Fang, X., Gu, H., 2017. Optimized throughput improvement of assembly flow line with digital twin online analytics, in: Robotics and Biomimetics (ROBIO), 2017 IEEE International Conference On. IEEE, pp. 1833–1837.

Sutharssan, T., Stoyanov, S., Bailey, C., Yin, C., 2015. Prognostic and health management for engineering systems: a review of the data-driven approach and algorithms. The Journal of Engineering 2015, 215–222.

Talkhestani, B.A., Jazdi, N., Schlögl, W., Weyrich, M., 2018. A concept in synchronization of virtual production system with real factory based on anchor-point method. Presented at the Procedia CIRP, pp. 13–17. https://doi.org/10.1016/j.procir.2017.12.168

Tan, L., Wang, N., 2010. Future internet: The internet of things, in: 2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE). IEEE, pp. V5-376.

Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., Sui, F., 2018a. Digital twin-driven product design, manufacturing and service with big data. The International Journal of Advanced Manufacturing Technology 94, 3563–3576.

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, 653–661.

Tao, F., Sui, F., Liu, A., Qi, Q., Zhang, M., Song, B., Guo, Z., Lu, S.C.-Y., Nee, A.Y.C., 2018b. Digital twin-driven product design framework. International Journal of Production Research 1–19.

Tao, F., Zhang, M., 2017. Digital twin shop-floor: a new shop-floor paradigm towards smart manufacturing. Ieee Access 5, 20418–20427.

Tao, F., Zhang, M., Liu, Y., Nee, A.Y.C., 2018c. Digital twin driven prognostics and health management for complex equipment. CIRP Annals.

Tidriri, K., Chatti, N., Verron, S., Tiplica, T., 2016. Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring: A review of researches and future challenges. Annual Reviews in Control 42, 63–81.

Trauer, J., Schweigert-Recksiek, S., Engel, C., Spreitzer, K., Zimmermann, M., 2020. What is a Digital Twin?–Definitions and Insights from an Industrial Case Study in Technical Product Development, in: Proceedings of the Design Society: DESIGN Conference. Cambridge University Press, pp. 757–766.

Tuegel, E., 2012. The airframe digital twin: some challenges to realization, in: 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA. p. 1812.

Tuegel, E.J., Ingraffea, A.R., Eason, T.G., Spottswood, S.M., 2011. Reengineering aircraft structural life prediction using a digital twin. International Journal of Aerospace Engineering 2011.

Uhlemann, T.H.-J., Lehmann, C., Steinhilper, R., 2017. The digital twin: Realizing the cyber-physical production system for industry 4.0. Procedia Cirp 61, 335–340.

Um, J., Weyer, S., Quint, F., 2017. Plug-and-Simulate within Modular Assembly Line enabled by Digital Twins and the use of AutomationML. IFAC-PapersOnLine 50, 15904–15909.

Vachálek, J., Bartalskỳ, L., Rovnỳ, O., Šišmišová, D., Morháč, M., Lokšík, M., 2017. The digital twin of an industrial production line within the industry 4.0 concept, in: Process Control (PC), 2017 21st International Conference On. IEEE, pp. 258–262.

van Otterlo, M., Wiering, M., 2012. Reinforcement learning and markov decision processes, in: Reinforcement Learning. Springer, pp. 3–42.

Wang, J., Ye, L., Gao, R.X., Li, C., Zhang, L., 2018. Digital Twin for rotating machinery fault diagnosis in smart manufacturing. International Journal of Production Research 0, 1–15. https://doi.org/10.1080/00207543.2018.1552032

Wang, X.V., Wang, L., 2018. Digital twin-based WEEE recycling, recovery and remanufacturing in the background of Industry 4.0. International Journal of Production Research 1–11.

Wärmefjord, K., Söderberg, R., Lindkvist, L., Lindau, B., Carlson, J.S., 2017. Inspection Data to Support a Digital Twin for Geometry Assurance, in: ASME 2017 International Mechanical Engineering Congress and Exposition. American Society of Mechanical Engineers, p. V002T02A101-V002T02A101.

Weber, C., Königsberger, J., Kassner, L., Mitschang, B., 2017. M2DDM–a maturity model for data-driven manufacturing. Procedia CIRP 63, 173–178.

Weyer, S., Meyer, T., Ohmer, M., Gorecky, D., Zühlke, D., 2016. Future modeling and simulation of CPS-based factories: an example from the automotive industry. IFAC-PapersOnLine 49, 97–102.

Wille, R., 2002. Why can concept lattices support knowledge discovery in databases? Journal of Experimental & Theoretical Artificial Intelligence 14, 81–92. https://doi.org/10.1080/09528130210164161

Xie, J., Wang, X., Yang, Z., Hao, S., 2019. Virtual monitoring method for hydraulic supports based on digital twin theory. Mining Technology 0, 1–11. https://doi.org/10.1080/25726668.2019.1569367

Xu, R., Wunsch, D.C., 2005. Survey of clustering algorithms.

Xu, Y., Sun, Y., Liu, X., Zheng, Y., 2019. A Digital-Twin-Assisted Fault Diagnosis Using Deep Transfer Learning. IEEE Access 7, 19990–19999.

Yao, X., Zhou, J., Lin, Y., Li, Y., Yu, H., Liu, Y., 2019. Smart manufacturing based on cyber-physical systems and beyond. Journal of Intelligent Manufacturing 30, 2805–2817.

Yun, S., Park, J.-H., Kim, W.-T., 2017. Data-centric middleware based digital twin platform for dependable cyber-physical systems, in: Ubiquitous and Future Networks (ICUFN), 2017 Ninth International Conference On. IEEE, pp. 922–926.

Zhang, G.P., 2000. Neural networks for classification: a survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 30, 451–462.

Zhang, H., Liu, Q., Chen, X., Zhang, D., Leng, J., 2017. A Digital Twin-Based Approach for Designing and Multi-Objective Optimization of Hollow Glass Production Line. IEEE Access 5, 26901–26911. https://doi.org/10.1109/ACCESS.2017.2766453

Zhang, H., Zhang, G., Yan, Q., 2018. Dynamic resource allocation optimization for digital twin-driven smart shopfloor, in: 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC). Presented at the 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), pp. 1–5. https://doi.org/10.1109/ICNSC.2018.8361283

Zhang, L., Zhou, L., Ren, L., Laili, Y., 2019. Modeling and simulation in intelligent manufacturing. Computers in Industry 112, 103123.

Zhang, M., Zuo, Y., Tao, F., 2018. Equipment energy consumption management in digital twin shop-floor: A framework and potential applications, in: 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC). Presented at the 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), pp. 1–5. https://doi.org/10.1109/ICNSC.2018.8361272

Zhang, Y., Ren, S., Liu, Y., Sakao, T., Huisingh, D., 2017. A framework for Big Data driven product life cycle management. Journal of Cleaner Production 159, 229–240.

Zhuang, C., Liu, J., Xiong, H., 2018. Digital twin-based smart production management and control framework for the complex product assembly shop-floor. The International Journal of Advanced Manufacturing Technology 96, 1149–1163.