applied sciences

Digital Twins in Industry

Edited by

A.Y.C. Nee and S.K. Ong

Printed Edition of the Special Issue Published in *Applied Sciences*

www.mdpi.com/journal/applsci

MDPI

Digital Twins in Industry

Digital Twins in Industry

Editors

A.Y.C. Nee

S.K. Ong

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade • Manchester • Tokyo • Cluj • Tianjin

Editors

A.Y.C. Nee
Mechanical Engineering
National University of Singapore
Singapore
Singapore

S.K. Ong
Mechanical Engineering
National University of Singapore
Singapore
Singapore

Editorial Office
MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Applied Sciences* (ISSN 2076-3417) (available at: www.mdpi.com/journal/applsci/special_issues/Digital_Twins_Industry).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-0365-1800-8 (Hbk)
ISBN 978-3-0365-1799-5 (PDF)

Contents

About the Editors

A.Y.C. Nee

A.Y.C. Nee is Professor Emeritus in Manufacturing Engineering at the National University of Singapore. His research interests include the use of AI, virtual and augmented reality in tool, die and precision engineering; sustainable product design and manufacturing; and digital twin applications. He is Fellow of CIRP, Fellow of SME, and Fellow of the Academy of Engineering Singapore.

S.K. Ong

S.K. Ong is an Associate Professor in Manufacturing Engineering at the National University of Singapore. Her research interests are virtual and augmented reality applications in manufacturing, ubiquitous manufacturing, sustainable product design and manufacturing, and applications of AI in manufacturing. She is a Fellow of CIRP, The International Academy for Production Engineering.

Preface to "Digital Twins in Industry"

Digital twin (DT) is an emerging and fast growing technology that provides a promising way to connect and integrate physical and virtual spaces seamlessly. In brief, a DT is a digital representation of a physical object or system. It has bi-directional communication capability with the physical twin through sensors and networks. DT is an evolution and integration of the various information-communication technologies (ICT) that have proliferated in the IT scene for the last two decades. It integrates internet of things (IoT), big data, cloud and edge storage, artificial intelligence of things (AIoT), augmented reality (AR), etc., to form a comprehensive communication network for controlling, monitoring, diagnosis and health inspection of equipment and facilities, traffic and transportation systems, buildings, etc.

DT has attracted much interest and enthusiasm from the academia as well as the industry. While academia has worked on algorithms and frameworks, industry will be the final implementer, as they can see the immediate benefits offered by DT technology. This book focuses on the industrial applications of DT technology, and it provides insights for the practitioners on how DTs can be successfully planned and implemented, as well as the desirable outcomes achieved.

This book contains 11 chapters covering a broad range of applications, in-depth review and integration of DT with other technologies such as AR and Industry 4.0.

In the chapter by Wärmefjord, et al., they discussed the barriers in the industry to be overcome before the use of DT for variation management and geometry assurance can be fully utilized. An extensive interview with engineers from eight different companies was conducted. They concluded that 3D models must be kept fully updated in order to maintain a robust digital thread.

The chapter by Sepasgozar advocated DT and web-based gaming technologies for online education. Not quite an industry application of DT as such, as it is more for educators. This is useful in view of COVID-19, as much of the face-to-face instruction has been virtual and online.

The chapter by Jacoby and Usländer emphasized the importance of interoperability by addressing the need to consolidate the various standards of DT and IoT. A classification scheme was created and applied to the standards to adopt serialization formats and network protocols to be used. An important issue, as this could lead to smooth and robust operations of DT and to overcome barriers of Industry 4.0.

An industrial application of DT was presented by Bambura et al. They implemented DT for engine block manufacturing processes. They constructed a DT consisting of three layers: physical, virtual and information-processing layers. Raw data were collected using programmable logic control (PLC) sensors. They concluded that even only partial results were presented, DT seems to be a prospective real-time optimization tool for the industry.

Another industrial application by Sierla et al. proposed a semi-automatic methodology for generating a DT of a brownfield plant, which is in the area of construction and urban development. Many procedures are required to construct a DT as outlined in the paper. The case study showed that only few manual edits were needed to the automatically generated simulation model.

In the chapter by Greco, a DT was used to set up models for monitoring the performance of manual work activities with near real-time feedback to support the decision-making process for improving working conditions. This is an interesting presentation of a human-centric DT for improving ergonomics and working conditions.

Autiosalo et al. presented an integrated DT for an overhead crane, providing service to machine designers and maintainers in their daily tasks. They showed that a good-quality Application Programming Interface (API) is a significant enabler for the development of DT, and advised traditional industrial companies to start building their own API portfolios.

In another industrial application, Pang et al. developed a DT and Digital Thread framework for an "Industry 4.0"shipyard. A new framework that combines the DT and Digital Thread was proposed for better management and to ensure continuity and traceability of information. The twin/thread framework encompasses specifications that include organizational architecture layout, security, user access, databases and hardware and software requirements.

The chapter by Pareja-Corcho et al. reported the development of simulation tools for the gerotor pumps. The paper is not a direct application of the DT, but is a virtual prototype that can be considered in the context of a DT tool. Future work is necessary to further integrate the physical pump with the software tool.

Agnusdei et al. presented an interesting chapter querying if DT technology supports safety management. The study analysed existing fields of applications of DTs for supporting safety management processes and provided a comprehensive bibliometric review to identify future trends between the DT approach and safety issues.

Carvalho and da Silva reported a rarely addressed area of DT-based systems in sustainability requirements. They conducted a meta systematic literature review and concluded that DTs across the product life cycle or the DT life cycle are not sufficiently studied. In addition, they mentioned in their research that it was not possible to find a paper discussing DTs with regards to environmental sustainability.

With the myriad of academic and industrial reports on DT development, this Special Issue could only represent a small fragment of the entire DT application scenario, not to forget the highly sophisticated commercial software that has been developed in recent years that are capable of handling large-scale and complex industrial systems.

DT is a promising technology, and its impact is yet to be fully realized in time to come.

This book would not have been possible without the dedicated contributions of the authors, reviewers and the editorial team of *Applied Sciences*.

The editors wish to congratulate all the authors and thank all the sterling support from the reviewers who have helped to refine the submissions.

Finally, we would like to place on record our sincere gratitude to Ms Jennifer Li, Associated Publisher *Applied Sciences*, who mooted this book and other assistant editors who relentlessly contributed their time and effort.

A.Y.C. Nee, S.K. Ong
Editors

applied sciences

Article

Digital Twin for Variation Management: A General Framework and Identification of Industrial Challenges Related to the Implementation

Kristina Wärmefjord [1,*], Rikard Söderberg [1], Benjamin Schleich [2] and Hua Wang [3]

1 Department of Industrial and Materials Sciences, Chalmers University of Technology, 41296 Gothenburg, Sweden; rikard.soderberg@chalmers.se
2 Department of Mechanical Engineering, Friedrich-Alexander-University Erlangen-Nürnberg (FAU), 91054 Erlangen, Germany; schleich@mfk.fau.de
3 School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China; huawang@sjtu.edu.cn
* Correspondence: kristina.warmefjord@chalmers.se

Received: 20 April 2020; Accepted: 8 May 2020; Published: 12 May 2020

Abstract: Digital twins have gained a lot of interest in recent years. This paper presents a survey among researchers and engineers with expertise in variation management confirming the interest of digital twins in this area. The survey shows, however, a gap between future research interest in academia and industry, identifying a larger need in industry. This indicates that there are some barriers in the industry to overcome before the benefits of a digital twin for variation management and geometry assurance can be fully capitalized on in an industrial context. To identify those barriers and challenges, an extensive interview study with engineers from eight different companies in the manufacturing sectors was accomplished. The analysis identifies industrial challenges in the areas of system-level, simulation working process, management issues, and education. One of the main challenges is to keep the 3D models fully updated, including keeping track of changes during the product development process and also feedback changes during full production to the development engineers. This is a part of what is called the digital thread, which is also addressed in this paper.

Keywords: digital twin; manufacturing; tolerancing; geometry assurance; digital thread

1. Introduction

Variation is an unavoidable element in mass-production. If not handled properly, variation can cause problems and significantly increase costs. Variation management is a broad term, relating to different methods to handle, and reduce the effects of, variation. In this paper, the focus is on geometrical variation. Problems related to geometrical variation usually constitute a significant part of the total cost for poor quality, sometimes up to 40% of the total cost for a manufacturing company in the form of delays, scrap, repair, rework, unsatisfied customers, and warranties [1,2].

Methods to reduce the effects of geometrical variation are sometimes referred to as geometry assurance. Additionally, terms such as tolerance analysis and tolerance management are used. Geometry assurance activities can be executed in different phases of the product realization cycle and cover areas such as the design of locating schemes/fixture layout, variation simulation, tolerance analysis, and inspection.

The digital revolution is expected to have a huge impact on the manufacturing industry in the coming years. Industry 4.0, increased digitalization, and access to an increased amount of data also open up for new methods and tools for geometry assurance, such as real-time process optimization using digital twins [3].

This paper aims at, from a standpoint in the current and future research areas and industrial needs, to identify the challenges, mainly in terms of processes and data flows, that need to be in place to fully capitalize on the possibilities offered by digital twins for zero-defect manufacturing.

A survey was distributed to experts in variation and tolerance management in academia and industry to identify the gap between research areas and needs, today and in the future. Not very surprisingly, this survey pointed out an increased demand for digital twins for improved geometrical quality. A deeper analysis using on-site interviews with over 40 geometry engineers from eight Swedish and Danish companies within the manufacturing sector was then conducted. The purpose of this more detailed study was to increase the understanding of the industrial needs and challenges related to the implementation of digital twins. Those results are the main contribution of the paper.

In the remaining part of the introduction, the basics of geometry assurance are described.

1.1. Geometry Assurance

Geometry assurance is a term describing a set of activities, all aiming to improve the geometrical quality of an assembled product. The geometrical quality affects both the esthetical and functional requirements of the final product. Usually, the maximum geometrical variation allowed in a certain dimension on the product is specified with a tolerance.

The geometry assurance activities start in the early design phases of the product realization process with tasks such as locating scheme optimization to secure a robust positioning system during the joining of parts and variation simulations to predict the variation in the critical dimensions of the final assembly. Inputs to the variation simulation are part geometries, information about joining and fixturing and inspection data, or suggested tolerances on the part level. The analysis is done iteratively to find a reasonable set of part tolerances, i.e., finding limits for maximum allowed variations on the part level at an acceptable cost. The simulation is sometimes based on the assumption that parts are rigid [4], and sometimes, the flexibility of, for example, sheet metal parts is included [5,6]. The latter one usually improves the simulation accuracy [5] and makes it possible to include clamping and joining forces, different material characteristics, joining processes, etc.

The variation simulation can be done in a standalone CAT (computer-aided tolerancing) software, such as RD&T [7], or integrated into the CAD system, as, for example, 3DCS [8]. Both setups are widely used industrially and have their pros and cons.

1.2. Scope of the Paper

This paper focuses on pinpointing future research and industrial needs within the area of variation management and geometry assurance. Digital twins are identified as an important area, and challenges, limitations, and requirements related to the implementation of digital twins for geometry assurance in the industry are identified and discussed.

In Section 2, the survey aiming to identify gaps between current and future research areas and needs is presented. Digital twins for geometry assurance are discussed in Section 3, and the interview setup and methods are described in Section 4. The findings are presented in Section 5, and challenges related to the implementation of a digital twin in an industrial context are highlighted in Section 6. This is followed by discussions and conclusions in Sections 7 and 8, respectively.

2. Materials and Methods

As a starting point for this work, a questionnaire was distributed among experienced researchers and engineers in the areas of tolerance analysis and geometry assurance. The purpose was to identify what research topics are perceived as most important today and in the future and, also, what the industrial status and needs are today and in the future.

2.1. Questions

A number of questions considering background variables were posed. Those questions gathered information about age, gender, years of experience in tolerance analysis, affiliation (industry or academia), and what continent the respondent lives on. This information, together with the answers to the four questions below, will be used for the analysis in this section:

1. Describe the area of tolerancing **research** of **today** by choosing a maximum of 4 keywords from below.
2. Describe the area of tolerancing **in the industry** of **today** by choosing a maximum of 4 keywords from below.
3. Describe important areas of **future** tolerancing **research** by choosing a maximum of 4 keywords from below.
4. Describe important areas of **future industrial development** by choosing a maximum of 4 keywords from below.

The keywords to choose between were the same for all four questions and can be seen below.

- **a. Tolerance specification**
- **b. Tolerance synthesis**
- **c. Tolerance analysis and evaluation**
- **d. Inspection and Metrology**
- **e. Simulation**
- **f. Digital twins**
- **g. Zero-defect manufacturing and self-compensation**
- **h. Digital thread (seamless digital flow of tolerance information)**
- **i. Tolerance management process**
- **j. New materials (composites, etc.)**
- **k. New manufacturing methods (AM, etc.)**
- **l. Perceived Quality**
- **m. Visualization**
- **n. Industry 4.0**
- **o. Robust Design**
- **p. Education—basic level**
- **q. Education—lifelong learning**
- **r. Other:**

2.2. Statistical Analysis and Method

This digital questionnaire was distributed among participants of the European Group of Research in Tolerancing (E-GRT) Biannual Seminar 2019 and other researchers and industrial engineers with an expertise in tolerance analysis and variation management. In total, 43 answers were collected. Information about the background variable responses can be seen in Table 1. For age and years of experience in tolerancing, the respondents could choose between different alternatives presented in the form of intervals. The midpoints of the intervals were used in the calculation of the average values presented in Table 1.

Table 1. Background information of respondents.

Number of Respondents	Gender	Average Age	Average Years of Experience in Tolerancing	Continent	Affiliation
43	Male: 37 Female: 6	40	9.8	Europe: 32 Asia: 11	Academia: 26 Industry: 17

The answers to the questions posed above were analyzed with the purpose to see the participants' opinions about current focus areas and future needs in research and the industry. To analyze the frequency differences among different answers, a Cochran Q test was employed. This statistical test is a nonparametric test for related categories with binary responses [9]. The related categories are the questions (a)–(r). This test can determine if there is a significant difference between the categories. The test statistic, for binary responses Y_{ij}, $i = 1, \ldots, n$, and $j = 1, \ldots, k$, is:

$$Q = \frac{(k-1)\left(kC - T^2\right)}{kT - R} \tag{1}$$

The variables are: n: number of respondents, and k: the number of questions.

$$C = \sum_{j=1}^{k}\left(\sum_{i=1}^{n} Y_{i,j}\right)^2$$

$$T = \sum_{i=1}^{n}\sum_{j=1}^{k} Y_{i,j}$$

$$R = \sum_{i=1}^{n}\left(\sum_{j=1}^{k} Y_{i,j}\right)^2$$

The Q statistic follows approximately a χ^2-distribution with $k - 1$ degrees of freedom. This approximation is considered valid if $n > 4$ and/or $nk > 24$. Here, $n = 43$, and $k = 18$.

The Cochran Q test is used to test whether there is a difference between any of the different categories. To point out what categories significantly differ from the other, a minimum required difference (MRD) test is used with Bonferroni adjustments to compensate for repeated tests [9]. The MRD is calculated as:

$$MRD = z_{adj}\sqrt{2\frac{kT - R}{n^2k(k-1)}} \tag{2}$$

where z_{adj} is the value of $(1 - \alpha_{adj}/2)$ in a standard normal distribution. The MRD is based on pairwise observations.

When comparing answers to the same question, with filtering based on the affiliation of the respondent, another test needs to be used. In this case, a standard z-test with pooled variance will be used [10]. This test compares the proportions, p_i and p_j, respectively, of positive answers to a certain question and test the hypothesis

$$H_0 : p_i = p_j$$

for two groups, i and j.

2.3. Results

Using the Cochran Q test previously described, it can be shown that there are significant differences between the alternatives (a)–(q) for all four questions: (1)–(4). The p-values for each question are all below 10^{-7}. For a deeper analysis of what categories (a)–(q) actually have a significant difference, the MRD is applied.

In Figure 1, the results from the survey are shown. Some observations can be made.

- In research today (left chart, blue bars), the most popular areas are simulation (e) and digital twins (f). For simulation (e), the MRD test shows that there is a significant difference ($\alpha = 0.05$) to all other categories except (c) and (f).
- In the industry today (left chart, red bars), the most popular areas are tolerance analysis and evaluation (c), followed by tolerance specification (a) and simulation (e). For (c), the MRD shows that there is a significant difference ($\alpha = 0.05$) to all other categories.
- For the future (right chart), digital twins (f) are considered highly important in both research and the industry. This is also true for simulation (e), new manufacturing methods (k), and Industry 4.0 (n). For research and digital twins (f)), the MRD shows that there is a significant difference ($\alpha = 0.05$) to all other categories except (e). For the industry and (f), the MRD test shows that there is a significant difference ($\alpha = 0.05$) to all other categories except (e), (g), (k), and (n).

To see the differences between the situation today and future needs, the differences between the answers are compared (see Figure 2). This illustrates the increased need of (f) (digital twins), g (zero-defect manufacturing), k (new manufacturing methods), and n (Industry 4.0), especially in the industry. For digital twins (f) it can be noted that the increased need for industrial development is not matched by an equally large increase in future research. Digital twins are already a hot research topic, as seen in the left part of Figure 1, and the challenge is the applicability of the digital twin framework in the industry. This will be the focus of a more in-depth interview study, presented in the following sections.

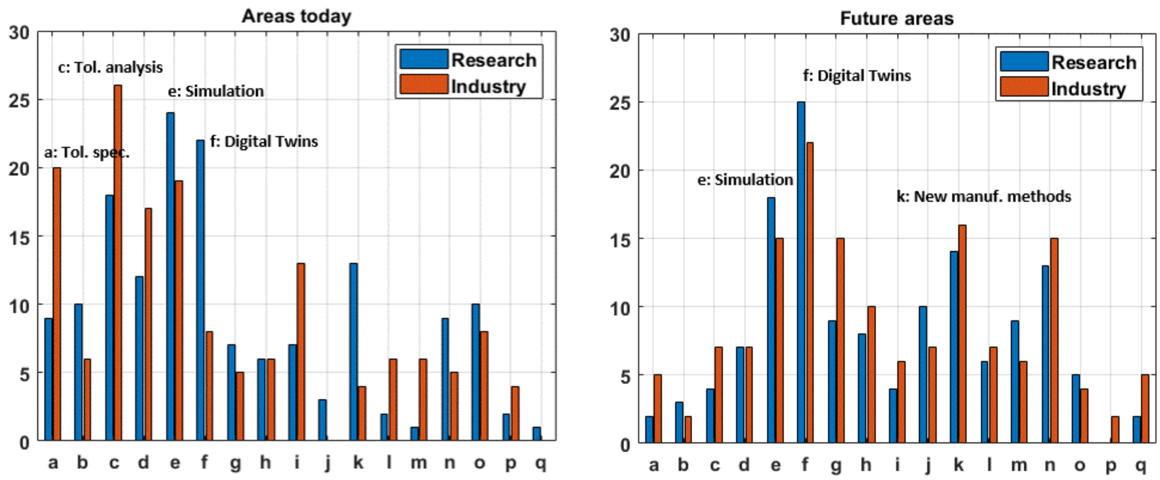


Figure 1. Responses to questions (1)–(4). The left part shows what areas are believed important today in research (blue) and the industry (red). The right part shows what areas are believed to be important in the future.

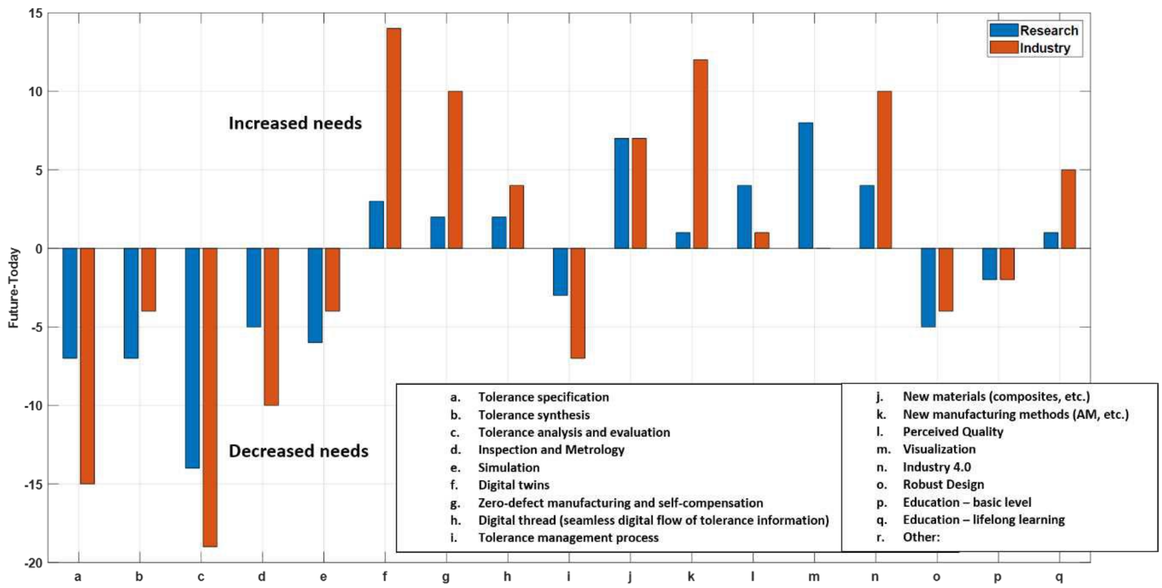


Figure 2. Gap between the future's and today's research and development needs in research (blue bars) and the industry (red bars).

Other interesting observations from the survey can be reached if the data is filtered based on the affiliation of the respondents (see Figure 3).

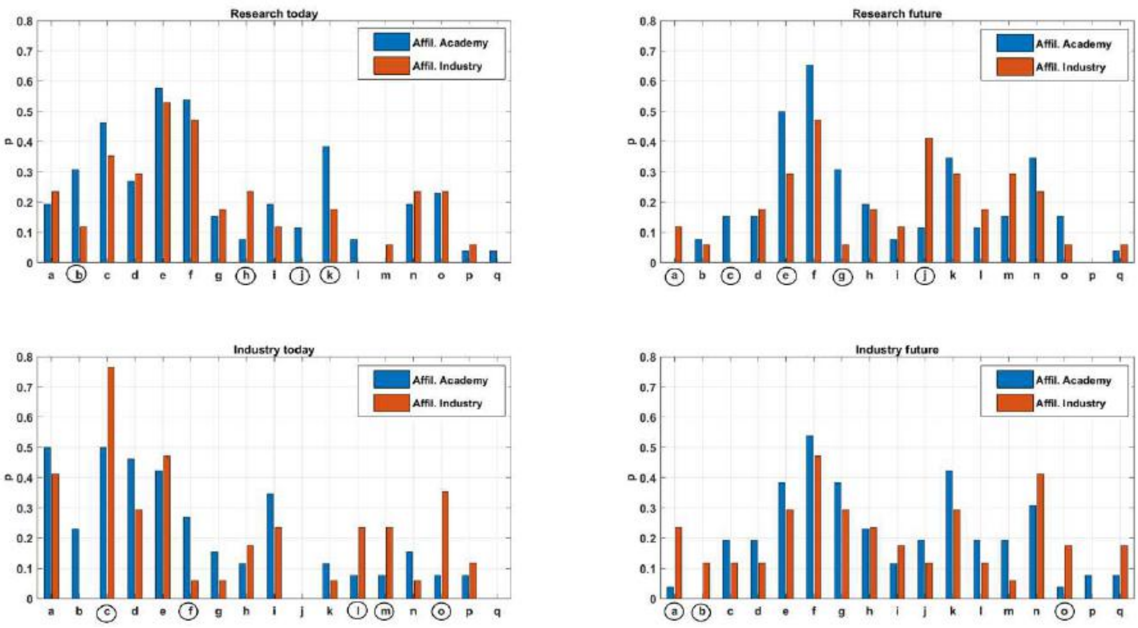


Figure 3. Responses filtered based on the affiliations of the respondents. Significant differences, using α = 0.10, are encircled.

By filtering data into groups, the sample size for each group decreases, and therefore, not all differences are statistically significant. The z-test introduced in Section 2.2 is used to test for significance. Significant differences, using α = 0.10, are encircled in Figure 3.

Some observations about the future:

- For research in the future, digital twins are highly valued in both academia and the industry. The significant differences are found in (c) (tolerance analysis and evaluation, (e) (simulation), and (g) (zero-defect manufacturing), which all are considered to be more important among employees in academia. On the other hand, (a) (tolerance specification) and (j) (new materials) are considered more important among industrialists than academics.
- For the industry in the future, the top priority is, again, digital twins. There is a higher level of agreement between people from academia and the industry in this area compared to research in the future, with significant differences only in (a) (tolerance specification), (b) (tolerances synthesis), and (o) (robust design).

In general, it can, based on the study, be concluded that digital twins are one of the most important areas within variation management. Today, it is most important in research, and in the future, it will become very important also in the industry, with an increased industrial need for development of the digital twins. In the next section, the meaning of a digital twin from a geometry assurance point of view is explained.

3. Digital Twin Framework for Geometry Assurance

The digitization of manufacturing is increasing. In the manufacturing business, 68% of the companies in a recent study stated that digital manufacturing is a top priority [11]. The number was even higher in countries such as India (94%) and China (87%). It is, however, also reported that the expected benefits are perhaps not yet fully reached. In the context of Industry 4.0, digitals twins are seen as one of the top ten technology strategic trends, according to Gartner Research [12], and various potential applications of digital twins along the product life-cycle have been identified [13,14].

A digital twin is a digital replica of a physical entity. In manufacturing, the terms digital model, digital shadow, and digital twin are used to indicate different levels of data exchange [15]. A digital model is a digital representation of a physical part or assembly without data exchange, a digital shadow enables a one-way data flow between the physical and virtual representations of an object, while a digital twin allows for a bi-directional automatic data exchange between the physical and virtual representation. In this regard, different reference models for digital twins have been proposed, such as in [16], considering important properties, such as model scalability, interoperability, expansibility, and fidelity. Most likely, a family of digital twins is needed to achieve optimized products and production flow, where each twin has its own purpose. The digital twins must represent both the product and the production system. As stated in [17], multi-scale analyses and distributed decisions will be required for optimizing the activities on the shop floor, as well as processes, resources, and machines. In many cases, digital twins will be interlinked. For example, the digital twin for geometry assurance described below must be linked to twins controlling the assembly cell equipment.

For the geometry assurance application, a digital twin approach was presented by Söderberg et al. [3]. This approach is also adopted in this paper. It was suggested that the simulation models, used in the design phases, can be reused and fed with inspection data during full production to do real-time individualized optimization of the assembly process to reach high-quality products. In Figure 4, an overview of the idea can be seen.

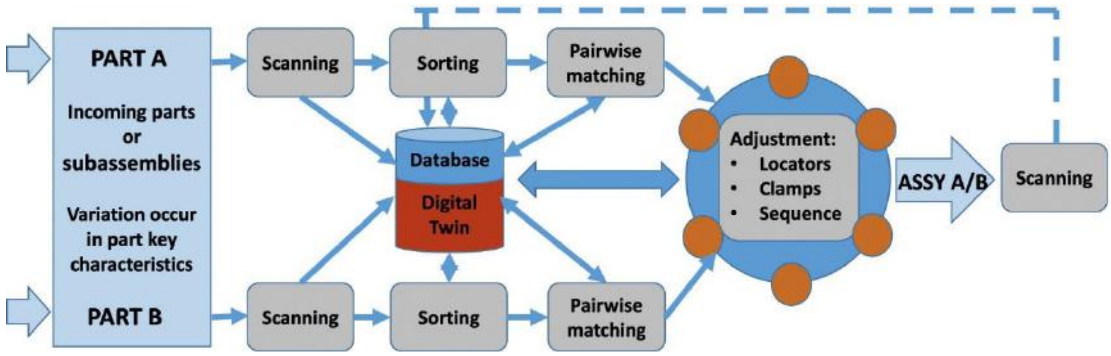


Figure 4. A digital twin for geometry assurance [3].

The idea is that two parts, A and B, are inspected using, for example, 3D scanning. This data is fed into the simulation model, which constitutes the kernel of the digital twin, to match individual parts over a batch of parts to minimize the geometrical deviation and variation of the assembly of A/B. This is also referred to as selective assembly [18]. Furthermore, the joining process can be optimized with respect to geometrical quality. Locators can be adjusted [19], and spot welding sequences can be optimized [20] for each individual assembly. By those suggested adjustments, the geometrical quality of the assembly can be improved without changing the tolerances of the parts, which is beneficial from a cost perspective.

With this approach, studies on several industrial cases have shown that the variation of the final subassembly can be reduced up to 50%, compared to when no adjustments are done [18–20].

With the labeling digital model, digital shadow, and digital twin mentioned at the beginning of this section, a digital twin for geometry assurance can be interpreted as:

1. a digital model, containing all nominal information about parts and assembly and joining processes. The model must also provide simulation capabilities.
2. input to the digital model about deviations from nominal values on the part and process levels.
3. output from the digital model to adjust the assembly and joining processes.

Those three pillars constitute together a digital twin.

3.1. *Digital Twin—The Digital Model*

As stated in [21], modeling and simulation are key aspects to implement a digital twin. To build the digital model, which is the core of the digital twin depicted in Figure 4, information about the joining and assembly process, part geometries, and other characteristics are needed. Using this information, a variation simulation model can be built. In the approach suggested in [3], the commercial software RD&T was used. This digital model is capable of relating deviation and/or variation on the part level to the deviation and/or variation on the assembly level.

Nominal dimensions, surface textures, material data, etc., which can constitute the input to build the digital model and complement the 3D model of a part, are sometimes referred to as product manufacturing information (PMI), according to ISO 16792. The use and development of PMI are, to some extent, driven by the fact that the 3D models successively are replacing 2D drawings [22]. This is beneficial not only due to its potential to reduce the amount of time dedicated to producing 2D drawings but also since the accuracy is supposed to increase.

Hedberg et al. [23] and other researchers at NIST (National Institute of Standards and Technology) discussed a digital model-based definition (MBD) for engineering tasks in manufacturing and inspection phases. This relates to the concept of a digital thread, which referrers to the digital information and data flow between different product realization phases. The digital thread provides important information to the digital model.

The 2017 ISO 1101 tolerancing standard allows for more precise tolerance definitions, which are independent of the viewing plane, and also supports the digital thread.

From a geometry assurance perspective, the digital thread must allow for a cohesive digital information flow between all activities related to product geometries, requirements, fixtures, and assembly cell layouts. This will be the focus of the interviews presented in Section 4.

3.2. Digital Twin—The Input Data

Scan data of the individual parts constitute the input to the digital model and can also be seen as a part of the digital thread. Given part deviations, adjustments adapted to a certain set of parts, A/B, can be predicted using the digital model. Of course, the scan data needs to be accurate and reliable. Scan data can be mapped to the nominal finite element meshes of the parts. Aspects of the scan data as the input to a digital twin for geometry assurance are discussed in [24] and are not the main focus of this paper.

3.3. Digital Twin—The Output Data

By feeding the digital model with inspection data, optimal locator adjustments and optimal spot-welding sequences can be determined. This constitutes the output from the digital model and is, of course, an important part of the digital twin. A suggestion of how to handle the data flow between the digital model and the assembly cell is presented in [25].

4. Interviews

To clarify the industrial state of the art and future needs regarding geometry assurance and digital twins, semi-structured interviews with over 40 engineers at eight Swedish and Danish companies in the manufacturing sector have been conducted.

The presented findings are focused on the information needed to build the digital model, using the terminology introduced in the previous section. However, to not limit the scope of the interviews and risk missing important aspects of the digital model and the simulation procedure, the whole geometry assurance loop is addressed. The importance of acknowledging interactions between different systems in a digital twin context is also highlighted in [26].

The engineers participating in the interviews work with geometry assurance or on a more general level with product development. Managers responsible for geometry assurance at the different companies were contacted, and they helped to select suitable interviewees. The interviewees were chosen based on their high competence and long experience in the area.

The base assumption of the interviews is that there exists a product development process containing concept, planning, and production phases. At most companies, this process is further detailed with gates and subphases [27,28].

A digital model, which will be the core of a digital twin, should be developed, reused, and updated in the different phases of the product development process in order to provide as much value as possible. Therefore, the respondents of the interviews were asked about the current situation regarding different gates and goals related to geometry assurance in the different phases (see Figure 5). They were also asked about the responsibilities of different roles, about input and output data for the different activities, and what (software) tools they use. The guide in Figure 5 was filled out together with an identical form reflecting what an ideal situation for the geometry assurance activities should look like in the future.

Current situation	Concept		Planning		Production	
What/why (gates and goals)?						
How—responsible?						
How—input data used?						
How—output data produced?						
How—what tools/processes?						

Figure 5. The interview guide.

As a part of the preparations for the interviews, the form, used as a guide during the interviews, was developed according to the recommended steps for a semi-structured interview described in [29]. Those steps are:

(1) identifying the prerequisites for the interviews.

In this step, the general procedure for the interviews was outlined. The ability to focus on issues being meaningful for the participants and also to allow diverse perceptions to be expressed were taken into consideration [30]. The companies participating in the study were contacted. All contacted companies accepted the interview invitation.

(2) retrieving and using previous knowledge.

Previous research in the area was investigated. The two interviewers had long (over 15 years) experience of research in the area.

(3) formulating the preliminary semi-structured interview guide.

The form illustrated in Figure 5 was developed. This form provided a framework for the interview and allowed easy movement from question to question, which was beneficial for the results [31] but, also, the possibility to dive deep into certain topics.

(4) pilot testing the interview guide.

To confirm the relevance of the developed interview guide, the questions were tested on an experienced engineer with many years of experience as a consultant within geometry assurance. Some small adjustments were made to make the questions clearer.

The interviews were held by two interviewers to reduced bias introduced during the interview. Before the interview started, its purpose and format were clarified. The respondents were free to elaborate on the questions, and follow-up questions were used to clarify answers.

After the interviews, the results were analyzed. One person took notes during the interviews and immediately afterward, those notes were clarified and validated with the other interviewer.

5. Results: The Geometry Assurance Process Today

In this section, the results from the interviews reflecting the present situation for the three main product realization phases are presented. An overview of the results can be found in Figure 6. In the figure, the different activities (middle) are listed together with the input required for the activity (top), as well as the output generated from the activity (bottom).

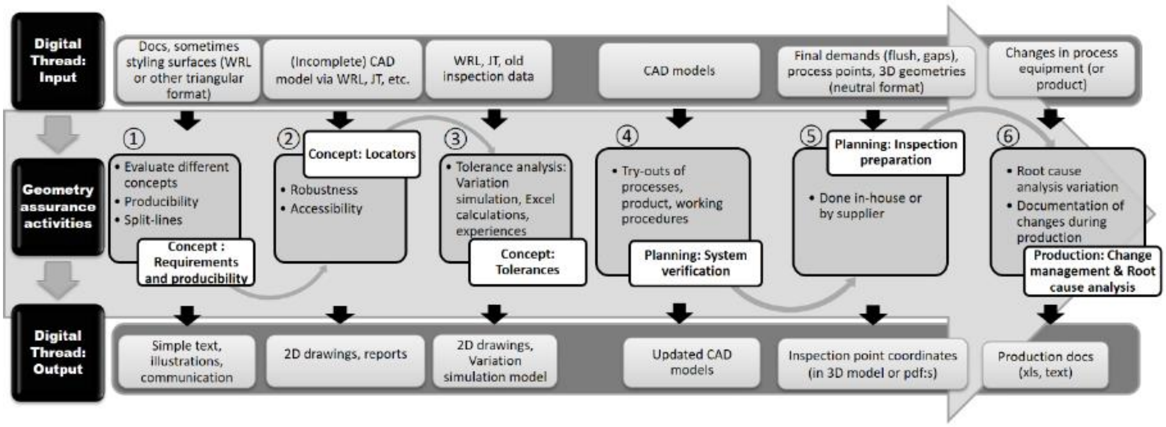


Figure 6. The input and output to the current digital thread for different geometry assurance activities.

5.1. The Concept Phase

For most companies in the study, the geometry assurance process starts with some kind of evaluation of concepts suggested by the design department. The suggested concepts are based on customer demands and other requirements and, to a large extent, also on experiences from previously produced models and projects.

The geometry engineers check if the suggested concept(s) can be built with reasonable geometrical quality (see ①, Figure 6). At this first stage, the evaluation is usually based on discussions, previous experiences, and sometimes, also on simple styling surfaces. The positions of the split-lines between parts affect both the visual sensitivity to variation and the level of geometrical variation in, for example, a weld between two parts [32].

The choice of locating scheme (②, Figure 6) is critical for the geometry assurance process. The locating scheme, used to fixate parts during assembly, must lock all degrees of freedom of a part and should be as robust as possible to variation. At this stage in the process, there usually exists a (not-so-detailed) CAD model of the part. This is exported from the CAD system in a neutral triangular format (often WRL or JT) and used in a variation simulation software to evaluate robustness. At the companies in this study, RD&T [7] is the most commonly used software. The suggested locating schemes are documented in reports and 2D drawings generated from the CAT tool. At one of the companies in the study, the produced results were linked to the 3D digital model.

The next main step is to define the tolerances on the parts (③, Figure 6). The goal is to find tolerances such that the final requirements on the product can be fulfilled but, at the same time, keep costs down. Again, triangular representations of the CAD model are used as the input, together with old inspection data (from similar concepts), if available. The analysis is done using variation simulation software to predict the final variation on the product (or subassembly) level or by simplified analyses in Excel or similar. The predicted variation is compared with the requirement on the product level, and the process is iterated until satisfactory results are obtained. Of course, experiences from previous projects are also important input. The output is the part tolerances described in a variation simulation model and/or on 2D drawings.

5.2. The Planning Phase

In the planning, or verification, phase, the interviews focused on system verification activities from a geometry assurance perspective and on inspection preparation.

The system verification (④, Figure 6) is often done through physical tests and try-outs. This is to complement and extend learnings from simulations. Of course, it would be desirable to reduce, or even exclude, the physical tests and rely only on simulations. This is a matter of "virtual trust" (i.e., trust in simulated results), a topic discussed in the following sections. Changes are done to increase producibility and quality. The changes can include minor changes of the geometries and changes of tolerances, fixtures/locating schemes, joining sequences, etc. After those changes, it is of course very important to update the CAD models and other information in the digital model and the digital thread.

One of the companies reported ongoing work to replace physical trimming, where a part is positioned in a checking fixture to control the shape of the flanges, with a virtual method based on the variation simulation model [33]. This is something that most companies want to do, but most of them are not there yet. Physical tests are expensive and should be kept to a minimum. At one of the bigger companies, it was reported that over 180 persons work with handling the most acute geometry-related problems exposed during system verification.

The inspection preparation (⑤, Figure 6) is an activity where the practices at the companies differ a lot. Some companies set inspection points according to the final requirements defined earlier and complement this with process points to monitor the process. At other companies, the chain from the final requirements to inspection points is not that clear. The inspection point preparation is done by the design engineer. Other aspects mentioned were that it is bad practice to do an inspection point preparation leading to points that cannot be evaluated/measured in a good way, that it is difficult to measure very flexible parts (due to effects from gravity), and that the number of inspection points often is too large and lead to difficulties in the evaluation and storage of data. Inspections are also associated with a cost, and it is desirable to keep the number of inspection points to a minimum. At the more advanced companies, 3D scanning is mixed with a CMM (Coordinate Measuring Machine) inspection. At other companies, gauges and other manual inspection tools are used.

For the digital thread, it is important to update the 3D models with inspection point coordinates. This is not always done. Quite often, an inspection point report is generated in PDF or similar, but the information is not fed into the CAD system.

5.3. The Full Production Phase

During full production, the geometry assurance activities are supposed to subside. Most of the geometry assurances should be preventive work in the early phases to make sure that the product and the processes are robust to variation. There are, however, always issues arising, and root causes of deviations and variations must be identified, and changes in the process should be documented (⑥, Figure 6).

Often, design changes done during full production are not brought back to the design engineers in a systematic way. There can be local systems, logs, and documentation, but it is not fully integrated into the design process. Therefore, it is a risk of repeating mistakes in future projects. Moreover, there is a disagreement between the virtual models and the physical world, and the generated inspection data will not fully reflect the virtual models. This has also been addressed in [34].

The importance of a good inspection database was mentioned at several companies. They used systems like QSYS [35] and CM4D [36] but were still missing some company-specific features. It can be concluded that the inspection database needs to be adapted to specific company requirements but, also, that different users have different needs and demands.

6. Results: Challenges Related to the Implementation of a Digital Twin

The future needs, identified during the interviews, that should be addressed to achieve a comprehensive digital twin for geometry assurance, are listed below. They can be divided into four different categories: system level, simulation working process, management issues, and education. The needs are further discussed in Section 7.

System-Level

(a) The major future needing identified at all the companies is a fully updated 3D digital model developing through all the phases in the product realization process. This is outlined in Figure 7. All activities and changes that affect the geometry of a part should also be brought back to the 3D digital model of the product, securing an always-updated and true geometrical representation of the actual produced part. This is today not the case at the companies in the study.

The 3D model should be updated when new information is available during the product realization process. Besides changes in geometry, locator positions, tolerances, and inspection point positions should also be linked to and updated in the 3D model.

Simulation Working Process

(b) Assembly models for variation simulations done in the early concept phases should be reused and refined during the planning phase. Today, usually new models are created, due to differences in modeling structures and working procedures. This is not only ineffective, but it can also lead to inaccurate information and misunderstandings due to model differences.
(c) The decomposition of final requirements to critical dimensions on the parts, done in the early concept phase, should be reused to a greater extent during inspection preparation in the planning phase.
(d) Simulation must be seen as one of the customers to inspection data, and the inspection should be adapted to this. This is necessary to provide high-quality input to the digital model and to achieve a complete digital twin.
(e) An individual digital twin of the product should be developed/kept for following-up purposes and for predicting maintenance.
(f) "Virtual trust" needs to improve.

Management Issues

(g) The design engineers need access to the production inspection data.
(h) There should be a generic project manning for all projects to secure that all aspects (including geometry assurance) are fully covered. This is not related to the digital twin concept but is included for completeness.
(i) Communication within the companies needs to improve. The design engineers are seldom willing to change their concepts to fit production capacity.
(j) People's willingness to change is sometimes low and needs to improve.

Education

(k) Education at universities should cover tolerancing and knowledge in geometry assurance to a greater extent. This finding is also supported by unstructured interviews with experts from German automotive suppliers, who reported that more attention should be paid to education in tolerancing and geometry assurance at universities. More particularly, the students should be introduced to the different steps of the geometry assurance process (as highlighted in previous sections) to fully understand the importance, interdependencies, and repercussions of tolerancing decisions on product quality and cost. Additionally, to cope with the challenges of a digitized design and manufacturing environment, students will need more knowledge and competencies regarding model-based definition workflows and the digital thread in geometry assurance in the future.

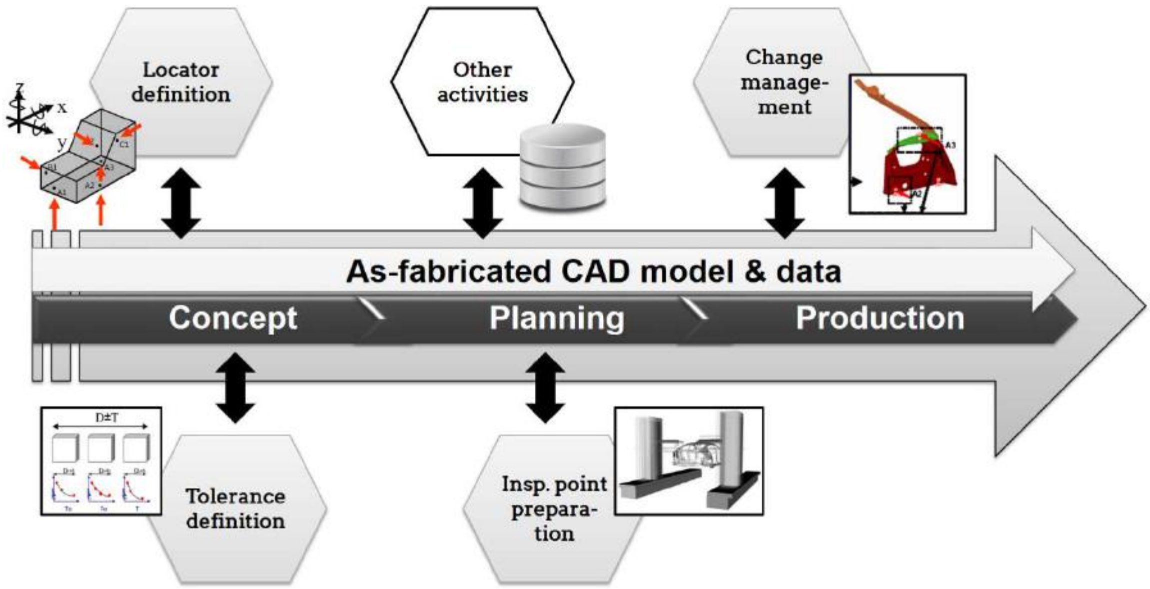


Figure 7. The CAD model needs to be updated with all new information from different geometry assurance activities (and other activities) to reflect the as-fabricated status.

7. Discussion

Among the identified needs listed in Section 6, the main need related to increase the use of digital twins and the digital thread supporting them is point (a) on the system level, the continuous update of the 3D models. If the 3D models only reflect the as-planned and not the as-fabricated product data, there is a huge risk for costly mistakes when using them in a digital twin concept. It is also difficult to learn from previous projects and improve new products without an updated 3D digital model. The reason for this lack of updated models might be practical, since some geometry assurance (and other) analyses are based on neutral triangular formats, such as WRL or JT, and the output might be 2D drawings or text files, not an updated version of the 3D model. Attempts to overcome this could be to use neutral formats—for example, STEP files—as both the input and output from variation simulation software. The STEP files can then be read into the CAD system again to update the geometry models. STEP files support PMI but not all GD&T concepts. An MBD approach where no 2D drawings are used, but the 3D models are the basis, removes some of the problems.

The CAT tool must be able to exchange information with the CAD system (and probably also the PDM/PLM system) regarding:

- geometry and material data,
- tolerances,

- locators (position/direction), and
- inspection points (position/direction) and inspection data.

The information can be exchanged directly or via the PDM system. It is also important to be able to quickly remesh the 3D geometries after changes [37].

The issues related to the simulation working process are also related to the development of the digital model. To improve "virtual trust", the simulation must be based on correct models and reflect reality. Updated 3D models and a clear digital thread are important aspects of this. Moreover, the inspection data used as the input to a digital twin must be of good quality. As stated, the simulation must be seen as a customer of the inspection data. For nonrigid parts, this means that the parts cannot be measured in an overconstrained position, which is usually the case. Instead, they should be measured by locking in only rigid body motions [38] to give a good representation of their actual shape. In that way, the variation simulation can predict the spring-back and the final shape of an assembly. Aspects related to how to choose the inspection points with maximum information content are discussed in [24,39]. Other factors affecting the accuracy of the simulation, and thereby the virtual trust, are listed in [5].

In the future, the interviewees state that a non-nominal digital twin of each individual product is desirable. This is also in line with the results of the survey presented in Section 2. If inspection data is linked to individual products via the digital thread, maintenance can be customized for each individual, which is especially beneficial for high-cost products.

Other aspects given in the list (g)–(k) are related to management and the education system. This is in line with the conclusions in [40]. Those aspects are not core parts of achieving a digital twin, but to use the full potential of a digital twin, they are probably necessary.

8. Conclusions

To conclude, a digital twin for geometry assurance shows great improvement potential. Examples have shown a reduction of variation on the assembly level with up to 50% compared to a standard joining and assembly process, without individual adjustments and optimization. However, there are barriers that the industry must overcome in order to fully capitalize on those potential improvements. Those barriers are mainly related to the lack of processes for updates and the sharing of the 3D models. Those models must be updated with all changes done during both the development and full production phases. To achieve a complete and reliable digital twin, inspection data must be of high quality and linked to the 3D models. Moreover, universities and research institutions have to elaborate on the underlying digital twin technologies and educational programs for engineers of the future to fully utilize the benefits of a digital twin.

Author Contributions: K.W.: conceptualization, methodology, formal analysis, investigation, and writing—original draft; R.S.: conceptualization, investigation, and writing—review; B.S.: investigation and writing—review; and H.W.: investigation and writing—review. All authors have read and agreed to the published version of the manuscript.

Funding: The work was carried out at Wingquist Laboratory and the Area of Advance Production at Chalmers within the project "Digital Tvilling", financed by Vinnova. This support is gratefully acknowledged.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Taisch, M. *The 2018 World Manufacturing Forum Report, Recommendations for the Future of Manufacturing*; World Manufacturing Forum: Cernobbio, Italy, 2018; ISBN 978-88-94386-1-0-3.
2. Thornton, A.C. *Variation Risk Management: Focusing Quality Improvements in Product Development and Production*; John Wiley & Sons: Hoboken, NJ, USA, 2004.

3. Söderberg, R.; Wärmefjord, K.; Carlson, J.S.; Lindkvist, L. Toward a Digital Twin for real-time geometry assurance in individualized production. *CIRP Ann.* **2017**, *66*, 137–140. [CrossRef]
4. Marziale, M.; Polini, W. Review of variational models for tolerance analysis of an assembly. *Proc. Inst. Mech. Eng. Part. B J. Eng. Manuf.* **2011**, *225*, 305–318. [CrossRef]
5. Wärmefjord, K.; Söderberg, R.; Lindau, B.; Lindkvist, L.; Lorin, S. Joining in nonrigid variation simulation. In *Computer-Aided Technologies-Applications in Engineering and Medicine*; IntechOpen: Rijeka, Croatia, 2016.
6. Falgarone, H.; Thiébaut, F.; Coloos, J.; Mathieu, L. Variation simulation during assembly of non-rigid components. Realistic assembly simulation with ANATOLEFLEX software. *Procedia CIRP* **2016**, *43*, 202–207. [CrossRef]
7. RD&T Technology. RD&T Webpage. Available online: http://rdnt.se/ (accessed on 8 April 2018).
8. DCS Webpage. Available online: http://www.3dcs.com/ (accessed on 8 March 2020).
9. Sheskin, D.J. *Handbook of Parametric and Nonparametric Statistical Procedures*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2003.
10. Larsen, R.J.; Marx, M.L. *An Introduction to Mathematical Statistics and Its Applications*; Prentice Hall: Upper Saddle River, NJ, USA, 2001; Volume 5.
11. McKinsey. *Report: Digital Manufacturing–Escaping Pilot Purgatory*; McKinsey: Boston, MA, USA, 2018.
12. Kerremans, M.; Burke, B.; Cearley, D.; Velosa, A. Top 10 Strategic Technology Trends for 2019: Digital Twins. Available online: https://www.gartner.com/en/documents/3904569/top-10-strategic-technology-trends-for-2019-digital-twin (accessed on 12 May 2020).
13. Schleich, B.; Dittrich, M.-A.; Clausmeyer, T.; Damgrave, R.; Erkoyuncu, J.A.; Haefner, B.; de Lange, J.; Plakhotnik, D.; Scheidel, W.; Wuest, T. Shifting value stream patterns along the product lifecycle with digital twins. *Procedia CIRP* **2019**, *86*, 3–11. [CrossRef]
14. Tao, F.; Zhang, M.; Nee, A.Y.C. *Digital Twin Driven Smart Manufacturing*; Academic Press: Cambridge, MA, USA, 2019.
15. Kritzinger, W.; Karner, M.; Traar, G.; Henjes, J.; Sihn, W. Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-Pap.* **2018**, *51*, 1016–1022. [CrossRef]
16. Schleich, B.; Anwer, N.; Mathieu, L.; Wartzack, S. Shaping the digital twin for design and production engineering. *CIRP Ann. Manuf. Technol.* **2017**, *66*, 141–144. [CrossRef]
17. Roy, R.B.; Mishra, D.; Pal, S.K.; Chakravarty, T.; Panda, S.; Chandra, M.G.; Pal, A.; Misra, P.; Chakravarty, D.; Misra, S. Digital twin: Current scenario and a case study on a manufacturing process. *Int. J. Adv. Manuf. Technol.* **2020**, *107*, 3691–3714. [CrossRef]
18. Rezaei Aderiani, A.; Wärmefjord, K.; Söderberg, R.; Lindkvist, L. Developing a selective assembly technique for sheet metal assemblies. *Int. J. Prod. Res.* **2019**, *57*, 7174–7188. [CrossRef]
19. Rezaei Aderiani, A.; Wärmefjord, K.; Söderberg, R.; Lindkvist, L. Individualizing locator adjustments of assembly fixtures using a digital twin. *J. Comput. Inf. Sci. Eng.* **2019**, *19*, 041019. [CrossRef]
20. Tabar, R.S.; Wärmefjord, K.; Söderberg, R. A new surrogate model–based method for individualized spot welding sequence optimization with respect to geometrical quality. *Int. J. Adv. Manuf. Technol.* **2020**, *106*, 2333–2346. [CrossRef]
21. Tao, F.; Zhang, H.; Liu, A.; Nee, A.Y. Digital twin in industry: State-of-the-art. *IEEE Trans. Ind. Inform.* **2018**, *15*, 2405–2415. [CrossRef]
22. Quintana, V.; Rivest, L.; Pellerin, R.; Venne, F.; Kheddouci, F. Will Model-based Definition replace engineering drawings throughout the product lifecycle? A global perspective from aerospace industry. *Comput. Ind.* **2010**, *61*, 497–508. [CrossRef]
23. Hedberg, T.; Lubell, J.; Fischer, L.; Maggiano, L.; Feeney, A.B. Testing the digital thread in support of model-based manufacturing and inspection. *J. Comput. Inf. Sci. Eng.* **2016**, *16*, 021001. [CrossRef] [PubMed]
24. Wärmefjord, K.; Söderberg, R.; Lindkvist, L.; Lindau, B.; Carlson, J.S. Inspection Data to Support a Digital Twin for Geometry Assurance. In Proceedings of the ASME 2017 International Mechanical Engineering Congress and Exposition, Tampa, FL, USA, 3–9 November 2017; p. V002T002A101.
25. Bohlin, R.; Hagmar, J.; Bengtsson, K.; Lindkvist, L.; Carlson, J.S.; Söderberg, R. Data flow and communication framework supporting digital twin for geometry assurance. In Proceedings of the ASME 2017 International Mechanical Engineering Congress and Exposition, Tampa, FL, USA, 3–9 November 2017.

26. Dahmen, U.; Rossmann, J. Experimentable digital twins for a modeling and simulation-based engineering approach. In Proceedings of the 2018 IEEE International Systems Engineering Symposium (ISSE), Rome, Italy, 1–3 October 2018; pp. 1–8.
27. Ulrich, K.T.; Eppinger, S. *Product Design and Development*; Tata McGraw-Hill Education: Noida, India, 2003.
28. Cooper, R.G. The drivers of success in new-product development. *Ind. Mark. Manag.* **2019**, *76*, 36–47. [CrossRef]
29. Kallio, H.; Pietilä, A.M.; Johnson, M.; Kangasniemi, M. Systematic methodological review: Developing a framework for a qualitative semi-structured interview guide. *J. Adv. Nurs.* **2016**, *72*, 2954–2965. [CrossRef]
30. Cridland, E.K.; Jones, S.C.; Caputi, P.; Magee, C.A. Qualitative research with families living with autism spectrum disorder: Recommendations for conducting semistructured interviews. *J. Intellect. Dev. Disabil.* **2015**, *40*, 78–91. [CrossRef]
31. Åstedt-Kurki, P.; Heikkinen, R.L. Two approaches to the study of experiences of health and old age: The thematic interview and the narrative method. *J. Adv. Nurs.* **1994**, *20*, 418–421. [CrossRef]
32. Wärmefjord, K.; Söderberg, R.; Lindkvist, L. Form Division for Welded Aero Components in Platform-Based Development. *J. Aerosp. Eng.* **2014**, *28*, 04014126. [CrossRef]
33. Lindau, B.; Rosenqvist, M.; Lindkvist, L.; Söderberg, R. Challenges moving from physical into virtual verification of sheet metal assemblies. In Proceedings of the ASME 2015 International Mechanical Engineering Congress and Exposition, Houston, TX, USA, 13–19 November 2015; pp. V02BT02A011–V002BT002A011.
34. Tao, F.; Sui, F.; Liu, A.; Qi, Q.; Zhang, M.; Song, B.; Guo, Z.; Lu, S.C.-Y.; Nee, A. Digital twin-driven product design framework. *Int. J. Prod. Res.* **2019**, *57*, 3935–3953. [CrossRef]
35. QSYS. Available online: https://swipx.com/apps/qsys-innovative-caq-softwaresystem/ (accessed on 17 March 2020).
36. CM4D. Available online: https://www.ats-global.com/products/ats-cm4d/ (accessed on 17 March 2020).
37. Cuillière, J.-C.; François, V.; Souaissa, K.; Benamara, A.; BelHadjSalah, H. Automatic comparison and remeshing applied to CAD model modification. *Comput.-Aided Des.* **2011**, *43*, 1545–1560.
38. Lindau, B.; Andersson, A.; Lindkvist, L.; Söderberg, R. Body in White Geometry Measurements of Non-Rigid Components: A Virtual Perspective. In Proceedings of the ASME 2012 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Chicago, IL, USA, 12–15 August 2012; pp. 497–505.
39. Wärmefjord, K.; Carlson, J.S.; Söderberg, R. A Measure of the Information Loss for Inspection Point Reduction. *J. Manuf. Sci. Eng.* **2009**, *131*, 051017. [CrossRef]
40. Schleich, B.; Wärmefjord, K.; Söderberg, R.; Wartzack, S. Geometrical Variations Management 4.0: Towards next Generation Geometry Assurance. *Procedia CIRP* **2018**, *75*, 3–10. [CrossRef]

applied sciences

MDPI

Article

Digital Twin and Web-Based Virtual Gaming Technologies for Online Education: A Case of Construction Management and Engineering

Samad M. E. Sepasgozar

Faculty of Built Environment, University of New South Wales Sydney, Sydney 2052, Australia;
Sepas@unsw.edu.au or samad.sepasgozar@gmail.com

Received: 8 June 2020; Accepted: 3 July 2020; Published: 7 July 2020

Abstract: Mixed reality is advancing exponentially in some innovative industries, including manufacturing and aerospace. However, advanced applications of these technologies in architecture, engineering, and construction (AEC) businesses remain nascent. While it is in demand, the use of these technologies in developing the AEC digital pedagogy and for improving professional competence have received little attention. This paper presents a set of five novel digital technologies utilising virtual and augmented reality and digital twin, which adds value to the literature by showing their usefulness in the delivery of construction courses. The project involved designing, developing, and implementing a construction augmented reality (AR), including Piling AR (PAR) and a virtual tunnel boring machine (VTBM) module. The PAR is a smartphone module that presents different elements of a building structure, the footing system, and required equipment for footing construction. VTBM is developed as a multiplayer and avatar-included module for experiencing mechanisms of a tunnel boring machine. The novelty of this project is that it developed innovative immersive construction modules, practices of implementing digital pedagogy, and presenting the capacity of virtual technologies for education. This paper is also highly valuable to educators since it shows how a set of simple to complex technologies can be used for teaching various courses from a distance, either in emergencies such as corona virus disease (COVID-19) or as a part of regular teaching. This paper is a step forward to designing future practices full of virtual education appropriate to the new generation of digitally savvy students.

Keywords: virtual reality; augmented reality; digital twin; 360 modules; YouTube; online App; construction; building; digital pedagogy; role play; e-learning; risk management

1. Introduction

Visiting a real construction site is not always possible due to site restrictions, the limited number of students permitted to enter a site, and, more recently, due to COVID-19. Virtual modules can be used for online education in architecture, engineering, and construction (AEC). They can also be applied to formative learning, flipped classroom [1], blended experimental teaching [2], and online teaching modes [3]. Previous studies investigated the feasibility of using virtual technology in education in different contexts, such as healthcare [4]. Recent studies intend to use virtual technology for education in construction and architecture such as Bashabsheh et al. [5], Wang et al. [6], Eiris Pereira and Gheisari [7], and Gao et al. [8]. However, the application of virtual technologies to show real physical practices in an immersive environment without using headsets for AEC education purposes has not been thoroughly investigated. There are complicated processes in construction, such as drilling and boring underground, which students have not experienced before, and traditional learning methods such as textbooks cannot easily deliver the knowledge. In contrast to text-based learning materials, there is a possibility of practicing in a simulated environment that allows students for correction and repetition to improve their skills with non-risk failures [9]. The purpose of this study is to present novel tools and online virtual applications to present the complicated processes of drilling, piling, and boring and an excavator digital twin to AEC students. The digital twin refers to the digital replica of a physical entity utilising the internet of things enabling two-way communications between them. The excavator digital twin and other education apps also address the deficiencies of traditional approaches in terms of promoting the engaging capabilities that allow students to be fully immersed in virtual space [6,10,11]. The purpose of these virtual apps is to enable large scale site visits, that will enable students to enter a virtual environment and learn a building case study or heavy equipment. The AEC courses may use project-based learning (PBL) approaches [12], so students are required to enhance their cooperation and collaboration skills [13]. Also, group project learning is recommended to all other educational disciplines at universities [14]. However, several severe challenges of PBL have not been adequately addressed for large classrooms. For example, the way instructors can measure each group member's contribution to their group assignment and give them immediate feedback in large classes will be much more difficult when it comes to the implementation of flipped learning methods or formative learning approaches.

The main research questions are:

(1) How can mixed reality and digital twins be applied in construction education?
(2) What virtual and augmented reality modules can provide a collaborative environment, and what factors may enhance students' engagement in construction education?
(3) What are the key values and advantages of the selected technologies in construction?

1.1. Significance and Advantages of Virtual Technology

The significance of virtual technologies is that it helps users to have an active experience rather than a passive learning experience and enhances their creativity [15,16]. Emerging technologies and virtual tools have caused a significant change in education methods, including construction education [17]. They shifted education and professional practices significantly away from the traditional individual theory-based lecturing to group PBL, similar to other practical disciplines. Project-based learning refers to learning from a specific construction project as a case study. Examples of group-based learning methods in engineering are problem-solving with open-ended solutions [18,19], hands-on projects [20], and team-oriented communications [21]. More recently, the concept of active learning and student engagement has had a significant impact on education design in practical courses [22], arguing that students learn more and are more prepared for their careers by actively applying the course materials. Some researchers recommended flipped classroom models for construction management [1,23]. However, this new highly lauded method has not yet received enough attention in the AEC, including the construction management discipline, particularly for large classrooms [1,24,25]. Also, there are not enough digital tools to support this teaching method. The problem is that the educator cannot take a large class of students into a construction site. This is particularly an important problem where a specific activity such as piling is not at the same time as the teaching period.

PBL is recommended to many educational disciplines at universities [14]. The PBL assessment is the core concern of many studies in different fields, including AEC, and students need to work collaboratively and enhance their social and cooperation skills. Table 1 presents the positive and negative experiences of students in doing PBL based on the literature.

Table 1. Overall advantages, experience, and barriers to group project assessments.

Students Experience	Description	References
Overall advantage	Encourage them to be more active and more engaged	Lee, et al. [26]
Positive experience	Reduce their workload and made their assignment more manageable	Finlay and Faulkner [27]
	Allow them to cover more materials with group members	Finlay and Faulkner [27]
	Their critical thinking skills would be improved through group conversation	Finlay and Faulkner [27]
Negative experience	Conversations were superficial and time-consuming	Bahar-Özvariş, et al. [28]; Herrmann [29]
	The group learning method took away lecture time where "facts" and direct answers would be provided	Lee, Ngampornchai, Trail-Constant, Abril and Srinivasan [26]
	Not productive because it takes time for the group to arrive at conclusions, mostly no results or consensus would be made	Lee, Ngampornchai, Trail-Constant, Abril and Srinivasan [26]
Barriers	Free-riders could be misidentified	Hall and Buzwell [30]
	Group members are not equally contributing to their assignment	Hassanien [31]; Hall and Buzwell [30]

1.2. Theoretical Factors

Virtual education utilizes a set of systems, including hardware and software, that provide an immersive environment or a "sensory illusion" to feel present in a different environment [32]. In virtual education, immersion, learners' perceptions, presence, and interactive activities are known as critical factors. The quality level of these factors, including immersion and interactivity, are related to technological attributes of the utilised technology such as digital images/videos, the display resolution, and other associated gadgets [32]. Lee et al. [33] suggests a set of factors that may increase users' satisfaction of a virtual reality module, including presence, motivation, virtual features, cognitive benefits, usability, reflective thinking. Some factors that are related to the technology acceptance model include usefulness and usability [34–36]. However, some other factors that are related to psychological aspects of learning, such as motivation, may enhance the effectiveness of learning and are usually considered in virtual learning tools. Lee et al. [33] also discuss that the cognitive factors are also important since they enhance understanding and memoriation of the learning subjects in the virtual environment. Memorisation increases the ability of students to recall events, facts, or definitions. Radianti et al. [32] discussed that 'immersion, presence, and interactivity' should be considered in technology design as core characteristics of the virtual module. From a psychological perspective, immersion refers to a state that the student feels isolated from the senses of being in the real world [32].

The literature discusses different factors that can be used for measuring students' experience of using newer technologies and predicting technology acceptance [37,38]. Nakarada-Kordic et al. [39] examined four principal measures of presence, such as feeling 'real', 'relaxed', 'comfortable', and 'anxious', while using the VR, before experiencing magnetic resonance imaging technology in medicine. They found that the VR experience improves participants' experiences before a potentially stressful use of the imaging examination. There are many published papers concerning sickness due to the use of VR applications. Kim et al. [40] developed a motion sickness metric for a successful VR implementation such as discomfort, fatigue, eye strain, difficulty in focus, headache, blurry vision, dizziness, head fullness, and vertigo.

The literature also presents the significance of tools or instructions to help students by enriching learning tasks, activities, and also with relevant experiences [41] (p. 3). However, this might be time-consuming and difficult to help students in large classes. A key component of providing active learning experiences in the classroom is that students need to be prepared for active engagement in the class [42] (p. 369). The flipped classroom is a change in the sequence in which activities are done by which students interact with the course materials [1]. In the flipped classroom, students preview the course materials before class so that they can do a part of their homework and other learning activities in class (workshop) [1]. As explained by Bliemel [43] (p. 113), core lesson materials can be available to students before class, including digital content and short videos. Preparing students before class or supporting them after class one by one is not easily possible in large classes. The purpose of this paper is to offer novel virtual technologies that potentially can be useful for online education and suitable to various class sizes and different teaching philosophies such as flipped classroom, blended learning, role-play learning, formative or summative approaches.

In particular, the current studies also investigate the positive and negative aspects of group project assessment and report that it is not possible to estimate the students' contribution to the main work accurately. However, some other studies provide some strategies to identify free-riders and/or assess other group members' contributions to the group project [44–46]. They offer different methods of peer evaluation or peer-assessments to assist lecturers in identifying the overall contribution of each group member. These methods are mainly based on a simple form asking students to give a mark of 100 to the group members based on their contribution to the group project. This form is an additional source to assess students' contributions to their group projects. However, there is not a reliable tool to detect biases and a valid measure to understand the level of each group member. Therefore, there is a need to address the mentioned problems by developing a novel technology-based model to evaluate group projects' individual marks accurately.

Learning management systems (LMSs) offer useful online tools for managing large classrooms [37,47]. Still, the current systems do not fully support a mix of educational approaches such as roleplay-based group project. In this paper, a set of tools, including a Group Wiki Project (GWiP) is offered as an essential tool for doing an online group project, as shown in Table 2. The GWiP is one of the necessary tools of Web 2.0 that provides spaces to write by students of the group in a web-based setting. This is a constructive activity and constitutes active learning in which students build an individual representation of their knowledge based on their peers' experiences [48].

Table 2. Characteristics of the collaborative learning platform.

Factor	Advantage	References
Sharing	Students can easily share their ideas and give immediate feedback to each other	Elmahadi and Osman [49]; Sonego, et al. [50], Gutiérrez-Braojos, et al. [51]
Collaborate and communicate	Allow a multi-player option enabling students to experience immersive, interactive and real-time meetings and working	Cooper, et al. [52] and Gutiérrez-Braojos, Montejo-Gamez, Marin-Jimenez and Campaña [51], Monahan et al. [53], Zheng, et al. [54]
Quick editing	Students can quickly edit their page	Biasutti and EL-Deghaidy [48]
Simple	The WiKi page is a simple mark-up scheme to set format (comparing HTML)	Dominic, et al. [55]
Convenience	Easy to link the page to other pages or sources (e.g., YouTube etc.)	Biasutti and EL-Deghaidy [48]
Maintainable	Easy to save the database, track all edits, logs and see the history and manage previous versions	Biasutti and EL-Deghaidy [48]
Open-source	Everyone who has access to WiKi can edit it	Dominic, Francis and Pilomenraj [55]

Männistö et al. [56] reviewed digital collaborative learning practices in nursing education and found out that this type of learning is beneficial since it contributes to knowledge construction and building on each student's interaction. At the same time, they suggest that instructors should provide more guidance to students and design a suitable pedagogical solution by using an appropriate tool for this type of practice [56]. There is a need to develop virtual tools for allowing the opportunity for collaborative learning or cooperative learning. In a collaborative learning process, students are not only working together to carry out their group projects, but they also need to work actively with their group mates and correct each other. Computer-supported collaborative learning systems should be developed further [57]. This type of collaborative learning is known as socio-constructivist [56]. Blau et al. [58] also discuss that this type of practice improves students' self-regulation skills and their learning proficiency by providing peer feedback while completing their project.

This paper first discusses a case study, including materials that should be taught in a construction subject; second, a variety of learning tools produced for online practice will be reviewed; and finally, an overall evaluation of students' feedback is discussed. Limitations and topics for future studies are also discussed in the discussion section. In this investigation, the forefront of contemporary advancements and innovation in AEC education was used to increase the authenticity of learning by being virtually present on site.

2. Research Method

In order to enhance students learning experience and improve virtual education in construction, this paper developed a set of virtual modules and discussed their applicability in construction. A practical construction course was used as a case study, and details of the course are presented [1]. First, a set of virtual modules have developed, as shown in Figure 1. Then, users tested and used them. The developed technologies were tested by a team of experts, including designers, technical programmers, educational developers, and students during the developing process, as well as before finalising the module. When the development process had been completed, a group of volunteer students was interviewed to learn their experience of using a virtual learning module. In order to improve education technologies, a scientific semi-structured interview was conducted among construction management students examining their experience and insights gained when using a selected virtual module. The semi-structured interview approach and content analysis are commonly used in the construction context [59–63]. The interviews were analysed manually by using the concept of thematic analysis. Students chosen as users of the modules were also selected to express their experience of using the modules. Their content on GWiP is used as learning evidence. With a focus on the interactive construction reality tour (iCRT) learning experience, the interviews interrogate the potential of the virtual module to support and elevate the students' engagement with the construction process through an immersive or interactive experience. This systematic data collection through interviews allows for a greater understanding of the student perspectives, learning processes, and adoption behaviour [64]. The users' feedback will enhance further development and adoption of VR tools and associated activity-level. Some topics that were asked to be discussed by the interview participants are as follows:

- How did a virtual module help you understand construction operations? For example, did you understand the link between excavators and trucks, piling rigs, and processes?
- If so, give an example of what have you learned or what was interesting to you in the module?
- What was your favourite part of the modules, and why?
- Do you think it made excavation operation process concepts clearer to you? And what subjects do you think complement or help assist in gaining a depth of knowledge?

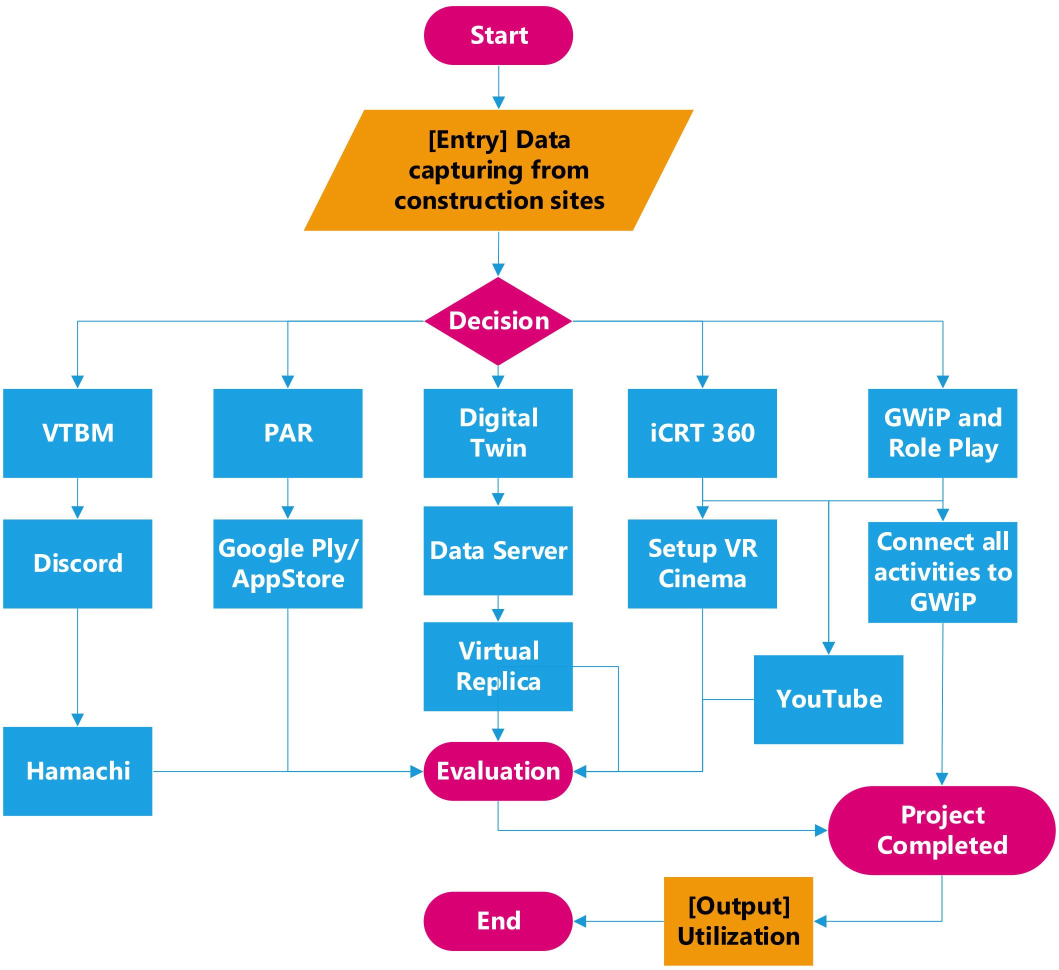


Figure 1. Flowchart of mixed reality modules developed for virtual education, including digital twin and online App. VTBM: Virtual Tunnel Boring Machine, PAR: Pilling Augmented Reality, iCRT: interactive reality tour, GWiP: Group WiKi Project.

The data collected from ten interviews were analysed in Sections 3 and 4. The students' feedback and their note on GWiP were used for identifying some key factors that may affect virtual technology acceptance.

Figure 1 shows the process, including data collection from construction projects to develop virtual learning modules, including VTBM, PAR, DT, and iCRT. These modules were created for students. GWiP is available on Moodle to students, and a short presentation of their work is available on YouTube. iCRT is available in a cylinder room, namely VR Cinema, and a simplified version is available on YouTube. VTBM is a virtual app that students will be given a link to download the file in conjunction with Discord and Hamachi to use it from home. PAR and DT are available on Google Play or App Store to students so they can download and use it anytime.

This section presents details of the selected courses, including objectives and topics which should be learned by students. The online modules, in this case, were developed to highlight principles that are difficult to recreate for students in the construction and engineering setting. The interactive resources were developed and used for the online delivery of relevant courses in different disciplines. This course was designed and improved by utilising different virtual technologies over five years from the time the course initially was created in 2016. The selected course is called Infrastructure and Industrial Construction (IIC), with the following details:

- First-year large foundational core courses. These students have not seen the entire process of underground activities.
- The course was designed based on collaborative learning and authenticity approaches in which unique virtual and online materials will be helpful;
- Large classes varying from 200 to 300 students, require developing digital materials for increasing the learning experience while being less expensive;
- Cover four learning outcomes to enhance students' ability to understand the processes and mechanisms of a variety of construction activities, as listed in Table 3.

Table 3. Selected course learning outcomes for the chosen course about infrastructure construction.

ID	Learning Statement	Related Assessment & Activities
1	Identify and analyse the management requirements, scopes, mechanisms, operational systems, performance and outcomes of a variety of contemporary national and international IIC projects	Quiz, Project and Final Exam
2	Analyse the construction processes and technologies used in infrastructure projects	Project and Final Exam
3	Critically analyse the strategic, tactical and operational challenges of IIC projects	Project
4	Demonstrate a range of professional characteristics required for IIC project personnel	Project

Table 3 shows the learning statements of the IIC course, which is a first-year core course with 300 students. This course was designed to extend students' knowledge of technologies, systems, and processes of industrial and infrastructure construction. This case study practised this approach by providing a portfolio of activities that includes students (i) forming their role playgroups, (ii) interacting with an industry guest lecturer, (iii) using the immersive environment (VR Cinema) to learn tacit knowledge about a construction case study, and (iv) learning from peers in group projects via GWiP. Arranging site visits for all students poses serious logistical challenges in terms of costs and personal safety considerations. These resources can transform the learning experience of students, very much in line with the objectives of the institutional strategies of improving the learning experience and the education literature. Selected topics covered in this course are shown in Table 4.

Table 4. Selected topics covered in the case chosen course.

Week	Topic
Week 1	Introduction to construction project organisations
Week 2	Site preparation and estimating earthmoving production
Week 3	Road construction and equipment
Week 4	Productivity, risk identification, and safety in practice
Week 5	Substructure, piling and tunneling construction
Week 6	Logistics, materials handling and mechanisation
Week 7	Trench shoring and temporary structures
Week 8	Superstructure and bridge construction
Week 9	Sustainable infrastructure and demolition technology

The GWiP was applied for the Industrial and Infrastructure construction course to give a chance to all students to do their group assignment based on their background and knowledge. The university strategy encourages lecturers to use innovative teaching methods. For a smooth transition between the traditional dominant face-to-face delivery model to the full online classroom model, a combination of different tools was designed and employed, as shown in Figure 1. To increase the authenticity of the modules, a set of engaging industry partners and world-class contractors were involved in the process of developing virtual modules caused in producing useful and valuable sources to students.

3. Technology Design

This section presents a set of tools and technologies to address the research questions of how mixed reality and digital twins can be applied in construction education and what virtual and augmented reality modules can provide a collaborative environment. These tools were developed for practical construction courses, as shown in Figure 2. Selected participants and stakeholders who are involved in this project include two leading construction contractors and consultants, the university estate management, the portfolio of the pro-vice-chancellor (education), a few instructors, and students from the Faculty of Built Environment, the Faculty of Engineering, and external technology vendors. Students at different education levels were involved in the project. In particular, two undergraduate students, two master's students, and one PhD student participated in the project of recording videos or preparing content for the VTBM module. These participants were involved in creating interactive teaching resources. These resources can be used in different courses such as construction informatics, digital construction, risk management, and practice-based courses, but were mainly used for designing the IIC pedagogy, so-called BLDG1021. Six novel interactive modules have been developed, as shown in Figure 2, and details are provided as follows.

3.1. Group Wiki Project and Role Play

The first module (refer to Figure 2) is an innovative group project online template, namely GWiP, which was designed for students to do their group projects together, where tutors and other instructors could monitor students' work in real-time. There is a lack of online tools available so far to show the students' progress in real-time transparently. In particular, measuring the student's contribution to their group project is always challenging, but the problem has been solved by the GWiP. All tutors and the instructor use GWiP to monitor students' progress weekly, give them relatively quick feedback, and increase the quality of their group project at the global level. Each group then presented their work (as a scenario base/role play based) and uploaded it on YouTube.

GWiP was perceived as useful technology helping students to prepare their group projects. Rogers [65] suggests that usefulness is one of the two critical factors of technology acceptance and can be used as a construct to predict a successful digital technology implementation. The advantage of GWiP was perceived as follows: students drafted their project gradually during the semester; instructors, including their tutor and lecturer, monitored their progress during the semester; the 'history' option on GWiP enabled the lecturer to check who contributed the group assignment more than others. GWiP's history page shows the number of words written by everyone with time.

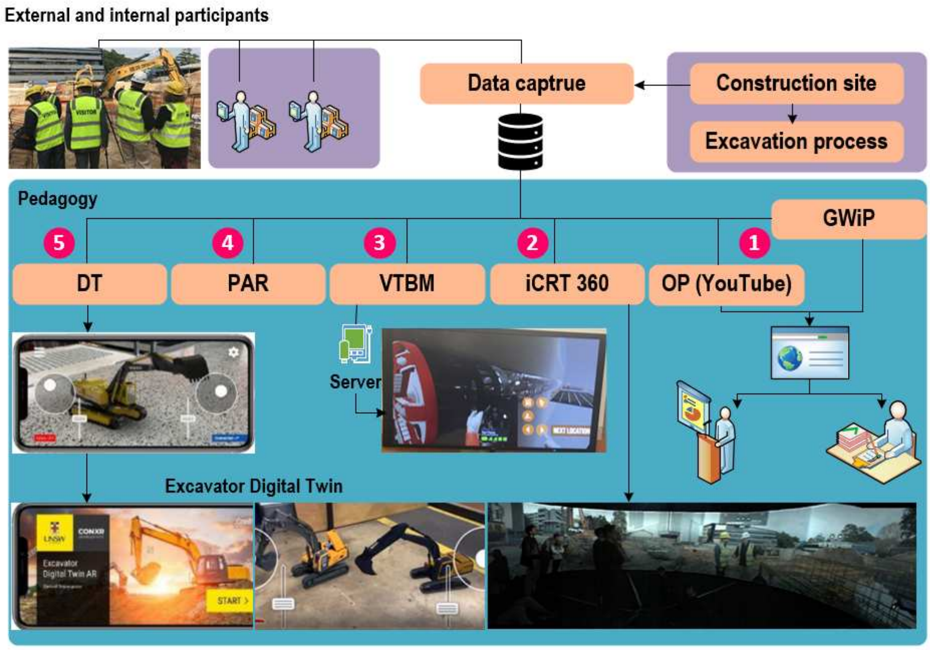


Figure 2. The flowchart of developed online modules, including screenshots of the applications, including Excavator Digital Twin.

GWiP increases the transparency of the teamwork, and students do not need to submit or print their projects at the end of the semester since all information will be saved automatically in their computer. For example, a total of 27 groups were formed by students in one semester. Then, 27 GWiP groups were created on WiKi Moodle by the lecturer each semester. Each student was asked to create a page within his GWiP. The group leaders also were encouraged to add extra WiKi pages for their group: 'executive summary,' 'project description,' and 'sharing ideas.'

The students were required to select and analyse a project (e.g., rail work, tunnel, highway, factory, and bridge) based on the relevant information, which they collected. They also were asked to describe their project (on GWiP) in a way that a person not familiar with the project could obtain a clear understanding of the project. The information was related to the course topics, including construction sites (e.g., location and accessibility), construction processes, project organisations, project monitoring, and their learning processes. Groups were encouraged to use theories and topics from the lecture throughout the semester and address them in their project. Based on GWiP, students were able to start their projects in two hours workshop running following the lecture and continue doing their projects any time per week in an online web-based setting.

Subsequently, each student analysed and discussed their chosen project on an individual online page. In this method, they do not use conventional word files, and instead, they described and reported their project on the GWiP page. The online pages allowed the lecturer and tutors to provide regular feedback on their drafts on GWiP. A page called 'Feedback on your Work' was created for each group, and the lecturer and grading tutors use this page to provide feedback to the students of each group. They were asked to identify features of their role, which presents something exciting, challenging, or unique about their role in their real project. They were provided with an example of people involved in the crane operation process, including crane supervisor, coordinator, operator, signaler, and slinger, as shown in Figure 3. It also shows that each group chose one project, and each individual took a specific role as examples of tasks. Figure 3 presents some of the groups and their workshop ID and the map page of their GWiP, their role, names, and ID numbers. In the first step, students take individual roles, then make the required online pages. They assessed the pages and received group and personal feedback on their work. The GWiP gives the possibility to the lecturer to assess students individually online, and at any time, so each individual within the group is assessable.

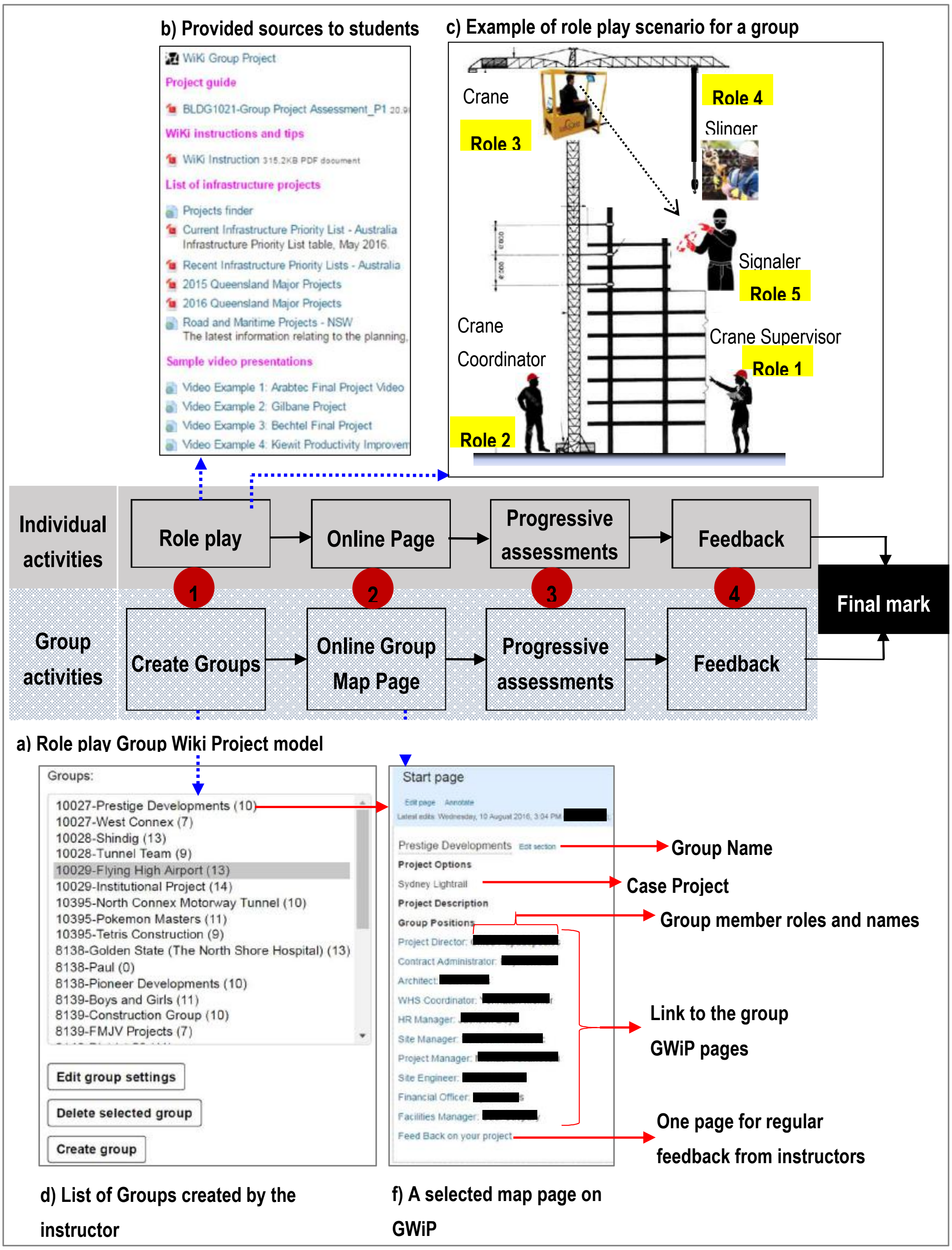


Figure 3. Role playgroup project model and list of groups and examples on GWiP. (**a**) Role play Group Wiki Project model; (**b**) Provided sources to students; (**c**) Example of role play scenario for a group; (**d**) List of Groups created by the instructor; (**f**) A selected map page on GWiP.

Then students presented the entire project in a video format and made it available on YouTube (search BLDG1021). Since the course was a large-sized class of over 200 students, applying to the role-play approach and presentation in class was not efficient, so students were asked to use computer visualisation aids and produce their videos and upload them on YouTube. This method helped the student to present the project they have completed during the term. This is in line with the project-based learning (PBL) approach [13]. The resources made by students are always available and helpful to next semester's students. Examples are shown in Figure 4.

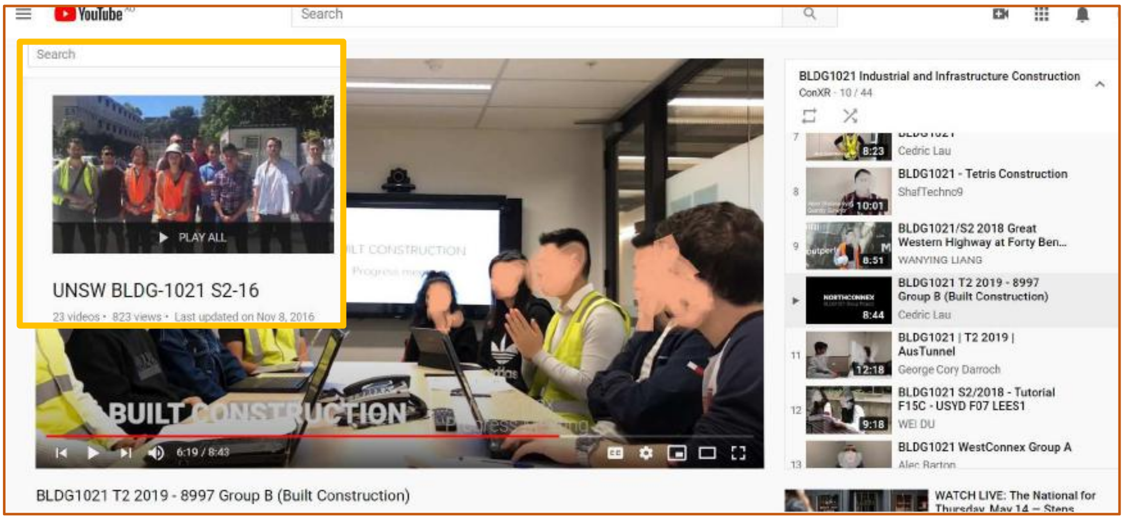


Figure 4. Students' presentations of their role-play practice uploaded on YouTube by themselves.

3.2. Interactive Construction Tour 360

The second online module (refer to Figure 2) is called iCRT 360, including videos recorded the real construction site of the Science and Engineering Building as a case study as shown in Figure 5. The content of recorded videos was combined with quizzes and interviews with the contractor (i.e., multiplex) on-site staff. The physical area of the immersive iCRT is a stereoscopic and interactive system. It consists of a cylindrical canvas which five high definition projectors visualise the image onto. These five projectors work collectively in sync allowing the 3-dimensional (3D) visualisation of movement to feel smooth and immersive.

Figure 5. Panning of an iCRT module sample module used for the case course [66,67] (See: https://www.youtube.com/channel/UCOzhGK8xOdoCc3Y9mJHnAnA/).

Additionally, motion tracking systems and stereo sound in the VR environment allowed students to interact with the projected modules allowing the opportunity for an immersive experience within an interactive VR environment. Added features in this format were the hotspots that detail specific parts of the construction process, including the equipment involved. Hotspots contained both short videos as well as photos to provide a more in-depth explanation of these processes and equipment. This feature allowed a degree of flexibility by varying emphasis on the description of different aspects of a construction site.

The iCRT uses the current technologies and cameras to capture rich information of complex construction activities for reproducing real practices meaning that students can visit noteworthy events during or any time after the event. The iCRT combines the current experience of Australian projects and the tacit knowledge of practitioners and uniquely involves them in digital education. The iCRT includes several modules covering underground and excavation activities in construction sites, as shown in Figure 6. For example, one of the modules focused on drilling and the pilling process. These activities are not visible, and students cannot see what is happening underground. Underground activities, excavation, and drilling are not safe for students to visit since there are many hazards in the areas these activities take place. Also, the entire process of piling, including excavation, inserting cages, concreting, may not be possible to carry out within visit times. Since these places are not safe, and the timing is also not under control, there is a need to visualise and create more interactive virtual resources to students.

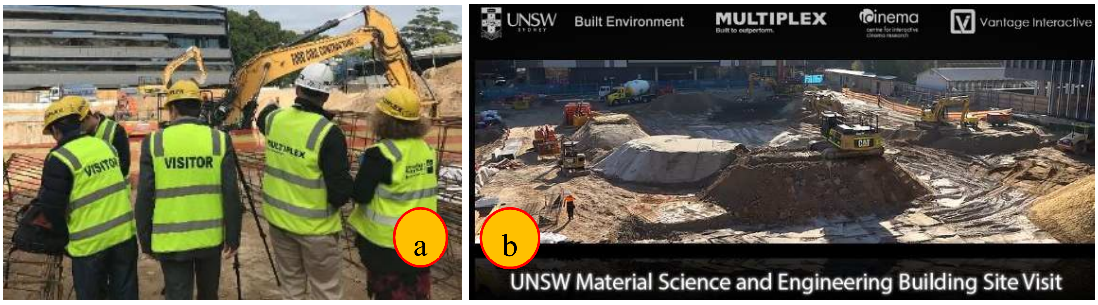


Figure 6. Leading a team of internal and external parties for developing the modules from fieldwork to digital lab production. (**a**) fieldwork for production planning in 2017; and (**b**) capturing 360 footages, including some patterns' logo involved in the production.

These resources aim to give the students a virtual experience of significant site and excavation works through cognitive learning in a workplace-based learning environment. Cognitive learning aims to teach learners the processes that experts use to handle complex tasks, situated within the context in which they would usually and naturally be carried out. It also aims to simulate the actual cognitive processes that have to be undertaken for a complicated task to be learned. A construction site, where students can find themselves surrounded by massive earthworks and equipment worth millions of dollars, can now be experienced first-hand through these new digital and virtual technologies, in a cost-effective, secure and efficient manner.

For more accessibility and further exposure (albeit at lower quality), the 360-degree VR videos were uploaded to a video delivery service that supports VR video, on YouTube. Students could view the VR video content through a desktop/mobile web browser, or through the Android/iOS YouTube apps. If students open these links up in Google Chrome, they can see details of the construction site or pilling process by zooming in/out. For example, if a student makes a mouse scroll, the user will get closer or farther away from the target. Students will be able to click and drag with a mouse to look in all directions of the construction site to explore all around the site.

Figures 7 and 8 show that the module is running using a server computer in the room behind the circular room. However, a tutor can control the module, including backward-forward, pushing the hotspots, and changes the modules using a tablet in front of the students. This is so students can become involved in running and managing the process of using the modules in groups of 10 to 20, which depends on the capacity of the cylindrical room as shown in Figure 9.

The virtual developed modules have responded to the disadvantage of digital education by dealing with the practical needs of the course. The virtual modules have the potential to change the students' learning attitude to be virtually present on site and increase the authenticity of learning by its virtual presence on-site and being able to see site managers operating in the modules. The modules allowed students to experience real practice, while savings on resources, transportation, time, and money compared with regular site visits made the virtual modules a much more sustainable proposition.

It also enabled practitioners to optimise their current processes to save more resources. The student's engagement was enhanced using these virtual modules, including iCRT, in the immersive environment. The simple version of my resources can also be used online remotely, which is precisely in-line with the UNSW 2025 Strategic Plan recommending the utilisation of "blended learning products with seamless integration of the physical and digital campuses". These products are critical for enhancing students' learning in construction since construction equipment is costly and project sites are often difficult to access.

Figure 7. The main computer is running the module in a separate control room.

Figure 8. The immersive environment used for the case course at the University of New South Wales (UNSW) (**right**), A selected Hotspot Site Layout Interview—Bulk excavation (**left**), the produced module, including the hot spots linked to rich information for each topic.

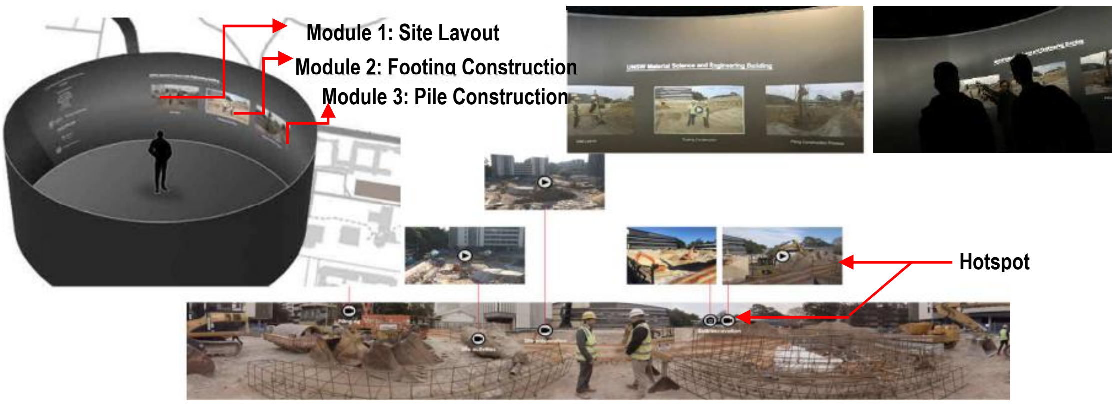


Figure 9. Three modules, including hotspots in the cylindrical theatre (VR Cinema). The construction site, including hotspots linked to 360 videos available to students.

3.3. Virtual Tunnel Boring Machine

The third online module (refer to Figure 2) is called VTBM. It is a game-based virtual environment allowing students to explore how a tunnel boring machine is working underground. This module provides a step by step process involving interactive virtual equipment, where students located in different areas inside or outside the university. They can enter virtually into the VTBM together (Figure 10). The number of students or groups of students allowed into the immersive virtual environment is not limited. Each tutor or student can invite up to ten students to join simultaneously and participate as group members to explore the immersive virtual environment together. The VTBM enables them to have the same experience with voice communication available and named avatars for all group members who can then see each other in the VTBM space underground. In VTBM, all students can walk individually through the virtual underground environment and explore different areas of the TBM located underground. In the virtual tour underground, they examine components and tasks relevant to TBM operations such as the cutter head, excavation chamber, mixing arm, bulkhead, screw conveyor, erector, tail skin, tunnel lining, hydraulic cylinders, and the backfilling process. The VTBM is based on a general 3DMax model of a tunnel and some images (e.g., 360 and 3D) collected from different activities so students will be able to explore more realistic underground movements and the TBM operation from any angle using a laptop or a HTC Vive headset.

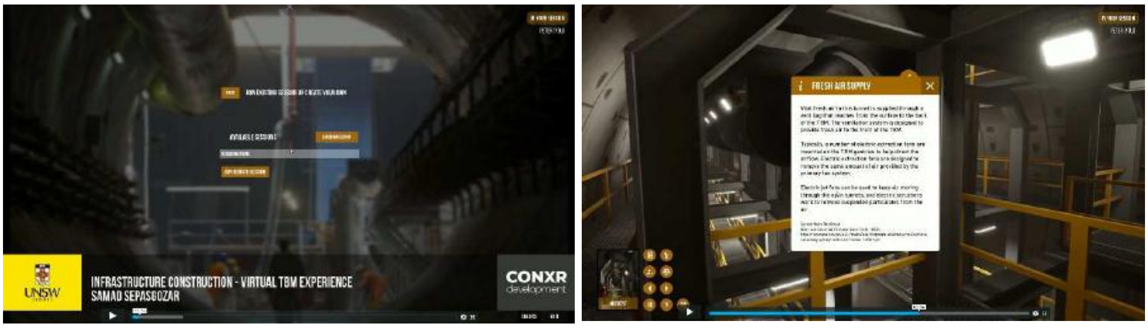


Figure 10. Screenshots from the login page (**left**), and one of the information hot spots on the VTBM (**right**). Several hotspots were embedded into the VTBM, referring to some interesting learning points in different parts of the module.

The VTBM is a multiplayer operation system module (See Figure 11), which can be used alongside Discord and Hamachi. Hamachi is required when the tutor allows more than one remote student on a different network outside the university to connect as if they are on the same network. Hamachi is a separate application that creates a type of virtual network over the internet. Students can discuss all components while exploring the TBM by using their device's microphone since there is a voice communication option that allows students to communicate with their teammates using Discord. Discord is an optional tool and should be installed separately. This will enable tutors to invite students into channels (potentially one channel per class) and keep the voice conversation continuous and transparent before and after the networked TBM experience is in action.

Figure 11. Utilising the VTBM on a large screen in an open studying space at FBE (**left**) and experiencing the VTBM using both a monitor and an HTC Vive Pro headset with another user logging into the VTBM from a different place in a collaborative multiplayer manner (**right**).

In a VR environment, there are 10 ordered drop-in locations, including appropriate thumbnail images, which let students experience ten specific identified sites and read the content provided on the hotspots (See the menu in Figures 9 and 12). Among the hotspots, one represents fresh air, and another refers to an air leak incident describing a fault and its consequences. These hotspots can be useful for risk registration and risk analysis.

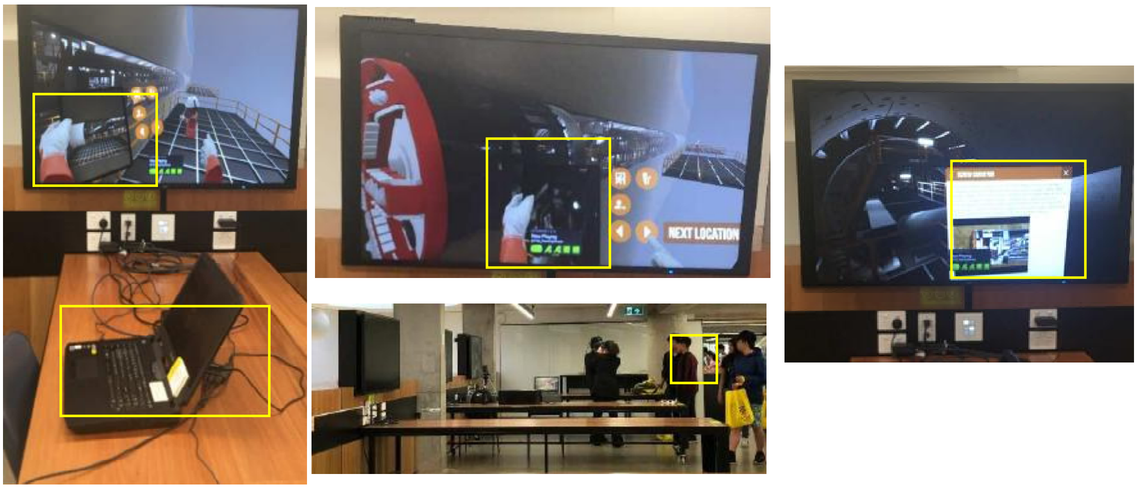


Figure 12. Screenshots from the menu and different sections.

3.4. PAR

The fourth online module (refer to Figure 2) is called 'FBE Piling AR' (PAR). FBE refers to the Faculty of Built Environment. The PAR is an interactive virtual environment that goes beyond the traditional pages of a textbook or PowerPoint and enables the foundation construction process to be explained in 4D (3D spatial models plus time). Students will have the ability to play through the animation of the developed typical example of the construction foundation piling work that supports the building. The PAR is an augmented reality app that is available on AppStore or Google Play to all global users. FBE can students download it, and with additional information provided in the course, they will experience the different processes of building construction, particularly piling methods and various types of piling failures. FBE's students can answer relevant quiz questions available on their course webpage on Moodle.

PAR was developed in two main versions: one version can be downloaded on smart devices (e.g., phones or tablets), and another version is available on Oculus headsets. An unlimited number of students can use PAR. However, if they want to experience the gaming environment as a group, up to ten students can enter into PAR simultaneously. Then they can see avatars in the headset representing their teammates. Avatars can see and ask each other questions when experiencing a different section of the App on the headset version. The avatars' appearance can be edited in Oculus Home, the virtual living room which a student launches into when the student puts on a Rift headset. They can save their experience in the augmented reality environment as a video or image formats, either to use the visual material later or to share/communicate their expertise with others/friends on Facebook.

PAR was designed for collaborative, interactive, and engaging practice and includes eight sections, as shown in Figure 1. It represents the construction process undertaken for a multi-story building and provides insights into structural foundation piles for students. The construction process covers site establishment, piling, and constructing the entire structure. The PAR experience is collaborative, i.e., all users see the same model in their VR headsets, allowing exploration and discussion as a group. The collaboration mechanism was enabled using Oculus Quest headsets (Figure 13) connected via a local Wi-Fi network. PAR offers both a mobile-based AR experience and an Oculus headset experience to learn from a construction case study project. In particular, students can observe different types of failure modes of foundation piles. Students across both platforms can view simplified structures and failure modes of the structural foundation elements (Figure 13).

The experience shows how the structural foundations of the building can fail due to the quality of the piling. The experience includes a foundation pile construction animation showing the entire construction process, including heavy equipment such as a drilling rig and excavator in different sections of the virtual animation. This animation gave the student the ability to explore the model section by section, and students were provided with additional information via hotspots around the model. The PAR model is based on 3D models of the Materials Science & Engineering building and the Kensington campus. The BIM file of that building, campus geographic information system (GIS) data [68] and light detection and ranging (lidar) data providing building height in a 3D context [69,70], were valuable in illustrating the built environment based on a real building. The BIM data available from the case study was used for representing different elements of the building from a pile to a completed building facade. PAR allowed students to look at the structural 'anatomy' of the selected structure (Figure 14) via a mobile device's AR interface or a VR headset (Oculus Quest).

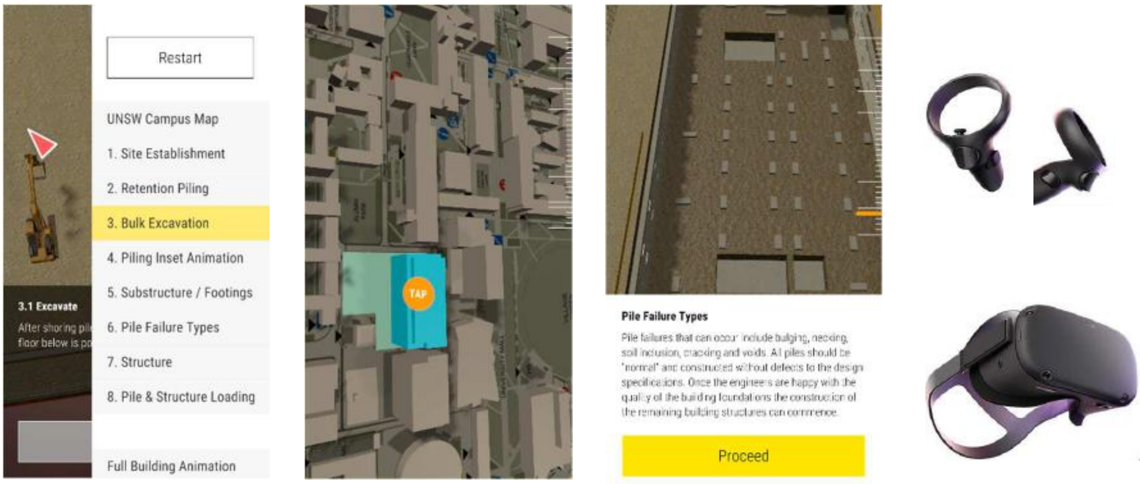


Figure 13. VR hardware for the collaborative model experience—the Oculus Quest.

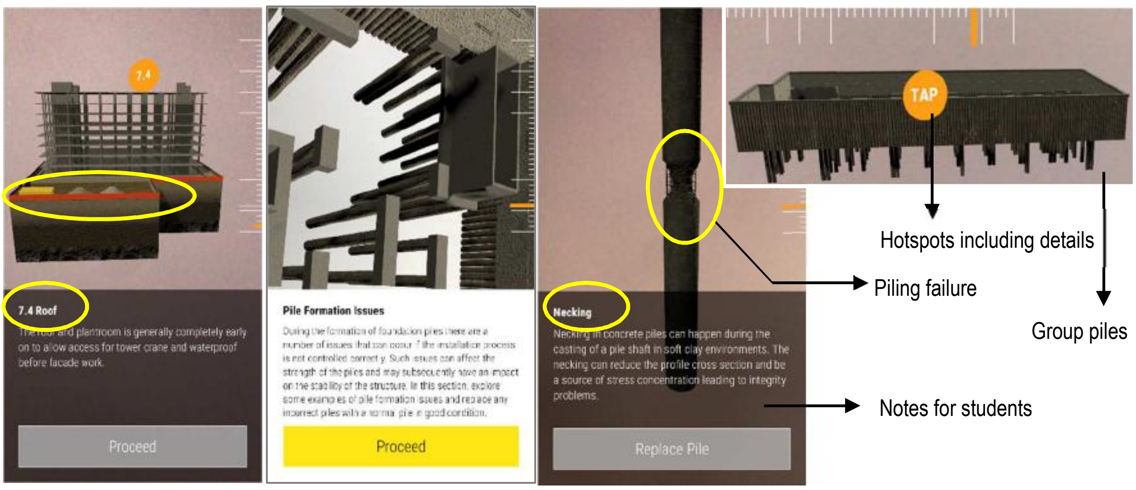


Figure 14. The collaborative model experience is presenting the entire structure and piling systems.

There is an introductory screen in the application to guide the students in how to complete the experience. If there are questions about the visualisation experience, students can ask the instructor or post their queries into a forum created on a learning management system (LMS) such as Moodle. In summary, the PAR offers (Figures 13 and 14):

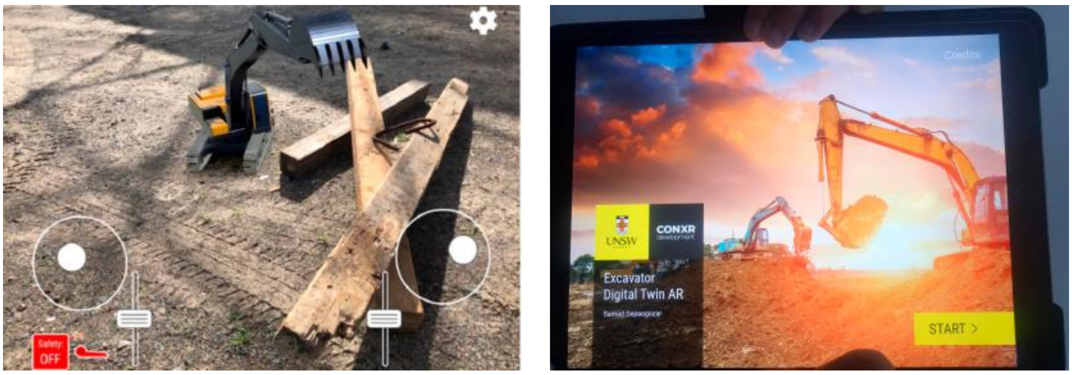
(i) An ability for students to look at the 'anatomy' of the structural foundations of the selected typical 3D building model;

(ii) Inclusion of a simple foundation construction sequence showing the different phases of a pile formation;

(iii) The opportunity for students to interact with information hot spots positioned around the model in 3D;
(iv) Views of the structures and typical failure modes of the foundation piles;
(v) AR phone app via the Apple or Android Apps if students use a compatible device (iPhone 6 or newer and AndroidS7 or newer), free to download for everybody;
(vi) Multiplayer VR experience of the 3D building model viewed within Oculus Quest headsets on the same local Wi-Fi network;
(vii) Visualisation and game experience on students' mobile devices (iPhone 6s or newer and Android S7 or newer); and
(viii) Visualisation on Oculus Quest headsets.

3.5. Digital Twin (DT)

The fifth online module (refer to Figure 2) is an excavator digital twin that is linked to a physical entity of an excavator. The DT module provides a virtual excavator so students can use it to learn different movements of the excavator. This is a step forward toward using a digital twin for education purposes. The connection between the digital twin and the physical twin required to be on campus, but using the digital version for simulation and students' practice is possible since it was developed for virtual education. This practice may change the educational approach for practice-based courses, as shown in Figure 15.



(**a**) (**b**)

Figure 15. The digital twin AR of the chosen excavator. (**a**) the AR excavator running in the field (https://www.globalconstructionreview.com/innovation/australian-academic-develops-digital-twin-digger/); (**b**) the digital model on iPad.

4. Interviews Results and Group WiKi Project Analysis

In order to address the research questions of what factors may enhance students' engagement and identify the advantages of virtual technologies, a group of students were invited to participate.

The GWiP documents of 204 students were also screened to identify how students played their roles and how they engaged in their projects. A group of ten students participated in the interviews to discuss their experience of using at least one of these technologies. Students were selected from the undergraduate program of construction management. Students were asked to focus on their roles, which have been taken in an infrastructure project and describe the challenges in their projects. Table 5 shows that they adopted their roles and learned about it. For example, one of a student described her role as a project director: "During my role as Project Director for the Sydney Light Rail Project thus far, many various obstacles have arisen. This is largely due to the size of the project and the high stakes the project holds for those involved in its construction and the impact the infrastructure project would have on the people of Sydney".

The result of the analysis of interviews about learning experience during the module design sessions shows the usefulness of virtual modules as well as limitations, which should be investigated in the future. The value and future research directions are discussed as following. Usefulness is a crucial factor in technology adoption and has been examined in the information systems [71] and construction over the years [64].

The participants support the claim that virtual modules, including the iCRT, provide an innovative immersive environment that brings real construction practices into universities to enable thousands of students from AEC to become familiar with state of the art in a no-risk environment [72,73]. Among these modules, the iCRT was recently experienced by over 1000 students. A student says: "It was more in-depth, and with the questions in the interactive exercises, we got to learn much more than just a YouTube video." Another student expressed the feeling that they visited a real construction site and learned from the visit: "I felt I was on an excursion on-site. It gave me knowledge and insight [into] what actually is involved in a construction site". Table 6 represents a summary of the student experience through the modules.

From the students' perspective, by "becoming immersed" in iCRT modules, a greater sense of "motivation and drive to seek further knowledge on key site elements" has been evident. This includes active engagement within quizzes of the quarry module, where trucks were used to underline the importance of Personal Protective Equipment. The modules provided to students have allowed the cohort to excel in "not just university, but also the workplace" by applying key aspects learned through the modules in the field as a student describes. A student expresses:

"When I first stepped on-site, it was a bit daunting. However, I quickly remembered key concepts of the site layout modules I learned in the VR labs and applied my knowledge to assist in site coordination and asserting the traffic management plan."

Several critical benefits have arisen from the application of virtual reality modules. Students can investigate how "productivity measures" are taken place through 360° cameras. An example of this practice includes the number of cycles an excavator takes to fill a truck from the bulk excavation. And through the alteration of crucial selections, the rate of productivity is evident, highlighting the emphasis of efficiency on site, correlating to both cost and time savings. Students have a real feel of the back of house cost and efficiency management, which can be applied on-site first hand in the field. Another benefit is the visualisation of machinery in practice and how events co-occur, highlighting the "concept of critical paths" as a student describes. "We were learning about Gantt charts and how activities have predecessors, and it was great to see how these activities were linked to successfully deliver a section of activities such as the laying of sheet piles." The concept of 'peering beyond the hoardings' highlights the unique experience of VR, producing "a depth of knowledge which cannot be gained unless on-site" one student states. Thus, the module including hotspots with additional and detailed information allowing for greater access to information, which in turn makes students more employable by increasing knowledge and skill base. Table 7 shows the key factors of virtual technologies from students' perspectives. The experience of an immersive environment is hugely enriching compared with just reading a textbook or looking at slides about the construction process.

Table 8 shows that virtual technologies allow for greater insight into construction methodology, expanding exposure to on-site practices. Activities included underground piling, excavating, and mining modules, focusing on the array of different aspects, including operations and risks involved. The virtual technologies directly complemented course content learned in lectures and tutorials where blasting of rock was seen, and shotcrete applied. This provides an insight into construction operations that occur mainly underground and cannot be seen unless on site. The virtual technologies encapsulate students through an innovative medium to transfer vital knowledge about construction, "safety, and scale of operations". This includes "visualising a piling rig's height in comparison to a human on-site with appropriate PPE" as one student describes. "It was great to see the whole operation of the footing system from the piling rigs drilling holes to the reinforcement being in place and finally in situ concrete being poured and capped off to create piles." The depth of knowledge can be seen as students have indicated the positive effect of VR in conveying construction methodologies and procedures.

Table 5. Summary of results based on the initial assessment and GWiP texts.

Expected Learning Outcomes	Student Learning
Understand scopes, components, and characteristics of national and international IIC projects	All groups selected one related project and described the characteristics of the project. For example: “*The main goal is to renew the CBD's urban landscape and mitigate congestion*” 8998 Group C, and “*This report provides an analysis of construction processes … through investigating the cross-London railway, Cross-rail … composing of 42km worth of tunnels …* ” Group 10395-North Connex
Understand the mechanisms, operation systems, performance and components of a variety of contemporary national and international IIC projects	The groups demonstrated that they understand the construction operation systems in their construction projects. For example, one of the groups (10029-Flying High Airport) discusses how the location of their tower crane should be determined. For example, “*The level by level approach requires 3 20t excavators to be lifted up to the top level with hammer attachments to demolish walls and slabs which would create a larger dust and vibration disturbance to the community.*” *8998 Group C*
Increase your ability to critically analyse IIC projects to explore strategic, tactical and operational challenges and suggest improvement scenarios	All groups tried to demonstrate their ability to analyse their project. For example, a group member explains: “ … *The successful management of Cross-rail is only made possible by various layers of project organisation focusing on their prescribed roles, with [a] key focus on deep construction understanding. This report finds that understanding construction processes will lead to beneficial outcomes, but obstacles that halt construction progress require active monitoring and key actions by management.*” Group 10395-North Connex
Enhance communication skills (i.e., professional writing and acting) and demonstrate how to work effectively in teams	The project is role-play based, and all of [the] students create videos and write on their Wiki pages, so they demonstrate their verbal and written communication skills.

Note: IIC stands for industrial and infrastructure construction.

Table 6. Features identified from the analysis of student experience from selected modules.

Key Factors and iCRT Features	Benefit	Students Experience and Feedback (Quotes)
Usefulness: Scanning the entire site	Clarity within the lessons regarding construction excavation processes	Step by step modules, on-site layout and footing construction assisted greatly.
Usefulness: Activity Analysis	Recognising the scale, size, and type of drilling and excavating equipment	*"As I had no on-site experience, it was extremely interesting to see the scale, size, and type of drilling and excavating equipment used on site. This included a man standing next to an immense piling rig, ultimately showcasing the mammoth operations undertaken on site."*
Rich sources of information: Hotspots (See Figures 8 and 10)	Hotspots provide me with rich content emphasising the significance of specific objects on site	Vividly indicates images of a piling rig, before, during and after shots to underline the concept of the activity to clarify construction techniques
Immersion: VR Cinema room/Innovation	An authentic learning environment that immerses me into the excavation and piling process despite being physically inside a classroom	*"The VR and 360 videos were unlike any tutorial or learning experience I had seen. It engaged me in an interesting medium, and I felt like I learned a lot more than regular tutorials and lectures from other subjects within the Construction Management degree."*
Enjoyment: Use of 3D Glasses and screens	Increases my interest in learning the subject	Innovation in technology has allowed for 3D glasses to be used within VR labs, creating a degree of realism. Through these innovative methods of teaching, we have students who have had an increase in interest in learning the subject.
Situated learning and Interrelated Concepts	Helps to relate previous knowledge to new information about excavation processes	Knowledge gained through other subjects in university complemented further information learned through the VR Cinema. In combining both, I was able to display my skills both at university and on-site effectively.

Table 7. Students' reflection and critical factors of the VR application based on semi-structured interviews.

Learning Outcome/Useful Features of the Module	Key Factor	Selected Quote
Required equipment for each construction task	Situated learning: types of machinery	*" … construction machinery is explicit [clear to us], but the use of it [applications] is difficult to know about it and understand .. how the machine works and the types of machinery … "* 8.08 HUST01
	How machinery work	*"Different [types of] machinery and the processes and .. how the machinery moves the dirt stuff … "*
Human resources and management	How to allocate people and machinery to a specific task	*"Consolidated my knowledge about how different machinery and different people that worked on site … "* 03.02
Deep understanding of processes and tasks	Detailed information for each component or task	*"The hot points … you can click on it, and you get more information and how interviewing people that work on the site … "* 01.51
Engaging students	Online feedback	*"Quizzes at the end of the video were interesting .. we could reflect on that some main points"* 01.38
Overview of the magnitude of operations	The scale of the construction site	*"It was great to see just how many vehicles and machinery are involved in bulk excavation and how many machines are operating around the site."*
Deep understanding of construction operations	Rich sources of information: Predecessors of construction processes	*"The twenty-minute modules allowed us to see the whole process of a piling rig from digging the hole to inserting the cage, in-situ concrete, and capping off; we were able to see the predecessors of the method."*
The conceptualisation of essential on-site practices	Labour and trades on-site	*"Being in the first year, we don't get to visit the construction site and have no construction experience under our belt, so being able to visualise how construction personnel works efficiently together was interesting."*
Operational safety awareness	Personal Protective Equipment	*"After seeing the scale of operations and sitting through a few modules, I was able to see what the potential impact of a downfall could look like after identifying a variety of risks on the construction site."*
Balanced practical and theory learning	Lectures covering theoretical concepts and VR tutorials allowing for hands-on experience	*"The VR and 360 videos complemented the theory we learned in the lectures … "*
Practical knowledge	Enjoyment	*" … as it was interesting to gain some practical experience."*
Workplace integration	Incorporating relevant principles	*"After looking back seven months, and working within the construction industry on-site, the information I had learned in the modules became relevant and useful."*

Table 8. Summary of functionalities and benefits of online teaching tools based on the instructor observation and the literature.

ID	Tool/Approach	Functionality	Benefit and Impact
1	GWiP (Refer Figures 3 and 4) and YouTube online group presentation	Creating a storytelling and project-based presentation and upload on YouTube by students.	Usefulness: Improve social skills [13]
		Drafting the project in an online platform, receive immediate feedback from instructors, and insert YouTube links and multimedia.	Engagement: A socio-constructivism practice [56]; knowledge construction and building on each student's interaction [56]; develop self-regulation skills and learning proficiency by providing peer feedback [58]; improve teamwork skills [13]
2	iCRT 360 (Refer Figures 5–9)	Virtual site visits and learn construction processes based on a case study.	Enjoyment: Digital storytelling [74], Learning by Seeing in 3D [75] in an immersive environment
3	VTBM (Refer Figures 10–12)	Virtual game-based environment to experience how a tunnel boring machine works.	Engagement: Engaging learning experience [4]; Learning by Seeing in 3D [75]
4	PAR (Refer Figures 13 and 14)	Virtual augmented reality to learn different building elements.	Enjoyment: Learning by Seeing in 3D [75]
5	DT	Digital twin of a real physical entity which students can interact with and learn how to operate an excavator.	Perceived usefulness: Deeper learning [74]

5. Discussion

This paper presents six innovative tools, including novel virtual digital applications, which specially developed for virtual teaching. The unique virtual technology implementation and students' feedback show that the use of online virtual tools for learning practical construction courses is feasible and useful. This paper describes how selected course content was designed and improved by utilising different virtual technologies over five years from the time the course was created in 2016. Students' satisfactions gradually have been increased, and the course was received firm quotations recently from students saying that they enjoyed and deeply learned the concepts which were not possible to learn from other resources. The use of virtual online technologies potentially enhances students' experience and engagement in practical construction courses. The immersive education modules have promising benefits to students, who are digital natives and known as tech-savvy [76]. These students have an inherent understanding of new online tools, smartphones, and digital devices. The concept of providing an authentic education space with interactive and engaging modules has seen an increase in the depth of knowledge and alignment with course content. Selected benefits were identified as the successful integration of theoretical concepts into practical experience in an authentic learning environment.

This investigation and modules created on virtual education consider the relevance of the online teaching approaches to allied academic disciplines, such as AEC as well as industry-led employee training. It thus speculates as to the online virtual advantages extending from education to economic deficiency and productivity. Visitors from other disciplines were inspired by the modules presented in this paper, including iCRT, and several leading global contractors and educators appealed to them from different universities.

The online versions of PAR and VTBM are significant educational solutions since they help students to obtain knowledge more rapidly and efficiently. The instructor can now spend more time in the online classroom, explaining other relevant concepts related to the visualised components in 3D and thus generate a more practical and complete lesson. This particular example of online practice also creates long-lasting knowledge because students are highly engaged in the space with the components of the building or the tunnelling machinery. Compared with a textbook or a lecture, students understand the practicalities and concepts more rapidly and with greater clarity. The produced modules provide an immersive sense of place when walking through a TBM or a building.

The main benefits of these modules are to improve the learning experience by taking students on a field trip in an immersive environment. In the virtual tour, students will experience a virtual representation of a real project and a simulation of a tunnel boring machine. This exposes students to a practical experience which helps them to understand how a tunnel boring machine works, and to learn the different structural elements of a sample building constructed by a leading construction contractor.

The results show that the online virtual tools are valuable and useful to instructors in construction to enhance students learning and assist them to retrieve their gained knowledge in the lecture in the real context. Future research is required to evaluate students' and tutors' perceptions of the model, using post-implementation qualitative data.

The value of this study is significantly high to AEC educators and scholars in this field who are from older generations teaching digital natives. This paper introduces digital practices to these educators who are accustomed to plan, work, and interact with others in a physical world rather than a virtual environment. The technologies presented in this paper can be considered as a subjective norm for or expectations of the digital natives. They have been using online apps and virtual game-based environments from their early years.

Restrictions on face-to-face teaching and learning due to the COVID-19 pandemic created more demand to use these types of advanced online multiplayer tools to replace construction site visits or laboratory experiences. These modules are designed to be useful for both individual learning and group interactions. They are much more comfortable for individuals to comprehend complex topics compared to textbook reading materials about construction processes. For example, students in a group of up to 10 people could see each other as avatars in a gaming environment and could explore different areas of the machine, talk about components, and interact with the gaming model. The students experienced the complex components or mechanisms of a machine in VTBM or elements of a building in PAR. The main practical implications of this study are listed as follows: This paper clarified the need and value for designing new online interactive tools for building construction; the practice of effective use of online tools in everyday teaching and learning was discussed; a greater awareness of online interactive tools was made, and relevant design and implementation issues were discussed, and finally a theoretical model was suggested to be examined as future investigation.

Several studies about the mixed reality show significant benefits for students, including improved learning effectiveness (76% more than traditional teaching), engagement, and motivation [77]. By giving students a self-paced interactive virtual learning simulation, students can repeat using the learning materials and experiments without additional costs. The online virtual modules have been designed based on the literature recommendation to include interactive resources such as embedded photos and videos and links to multiple-choice questions. The interactive elements provide different learning scenarios to allow construction students to practice safely and to transfer their knowledge to practice. Where some current commercialised modules simply use expensive digital technologies (e.g., oculus gears and goggles) applicable to small classes, the presented innovative resources were used for massive classes of students. It was not possible to provide oculus gears and goggles to 300 students, and also, it was not possible to take all of them into a construction site. The experimentation showed that the virtual and augmented reality modules gave students a chance to explore building construction processes. They could also experience a tunnel boring machine in operation, identify potential operational risks and hazards (e.g., foundations cracking or leaking issues), and discuss issues while experiencing the immersive environment.

As another example, GWiP is a novel real-time group project model. This model makes the main contributions to the field. First, the GWiP model encouraged students to discuss with piers as each of them takes a role similar to the real projects, and they gave them immediate feedback. Also, tutors could monitor their work in real-time and were able to provide them with feedback when they have any questions. This is in line with the previous recommendation in the literature. For example, Gómez-Pablos et al. [13] discussed that students need help and support for doing their tasks because sometimes they do not know how to work effectively with their group mates or peers. They suggested that their social skills need to be reinforced and improved while doing different tasks at the university [13]. The GWiP helped students to know each other, trusted to peers, and supported each other for constructively accomplishing their assigned works. The role of GWiP model was to allow the opportunity to improve students' interpersonal skills, thereby providing a digital collaboration platform for practising the required skills. The GWiP has also encouraged students to start working on their homework in the workshop under tutors' supervision and continued doing the project outside of class using an online platform under the instructors' supervisor. The GWiP model enabled instructors to examine student's contributions to the group project report in real-time during the semester. However, previous approaches rely on students' judgments that use a different scale of measure.

Based on the interviews, module development experimentation, and the literature, this paper suggests three main factors, including 'perceived usefulness', enjoyment, and engagement as three primary constructs of satisfaction. All these factors can contribute to modelling and predicting virtual technology acceptance. This model is suggested as a conceptual framework that can be examined in different contexts and can be considered as research hypothesis in future studies:

Research question 1: Technology acceptance modelling has a lengthy theoretical background and applied in a different context, but this paper contributes by presenting an extended model applicable to virtual education tools [64,71,78–80]. The article shows some factors that can be considered as measures of "usefulness", which is one of the critical constructs of the technology acceptance model [78,81]. This paper also suggests that 'usefulness' of the virtual technology refers to students' experience of social presence, the possibility of using a rich source of information, and situated learning, which all help students to comprehend detailed practical information of operation process as shown in Figure 16.

This is useful since otherwise, it is not possible to obtain the information without involving in a project as a cadet or intern. Figure 16 shows a list of factors identified in interviews and the literature that can be used for modelling and predicting virtual technology acceptance by participants. The proposed model can be examined for different technologies such as DT, VR or web-based virtual technology. The technology can be used for different subjects covering construction operation, risk analysis, safety, and construction informatics.

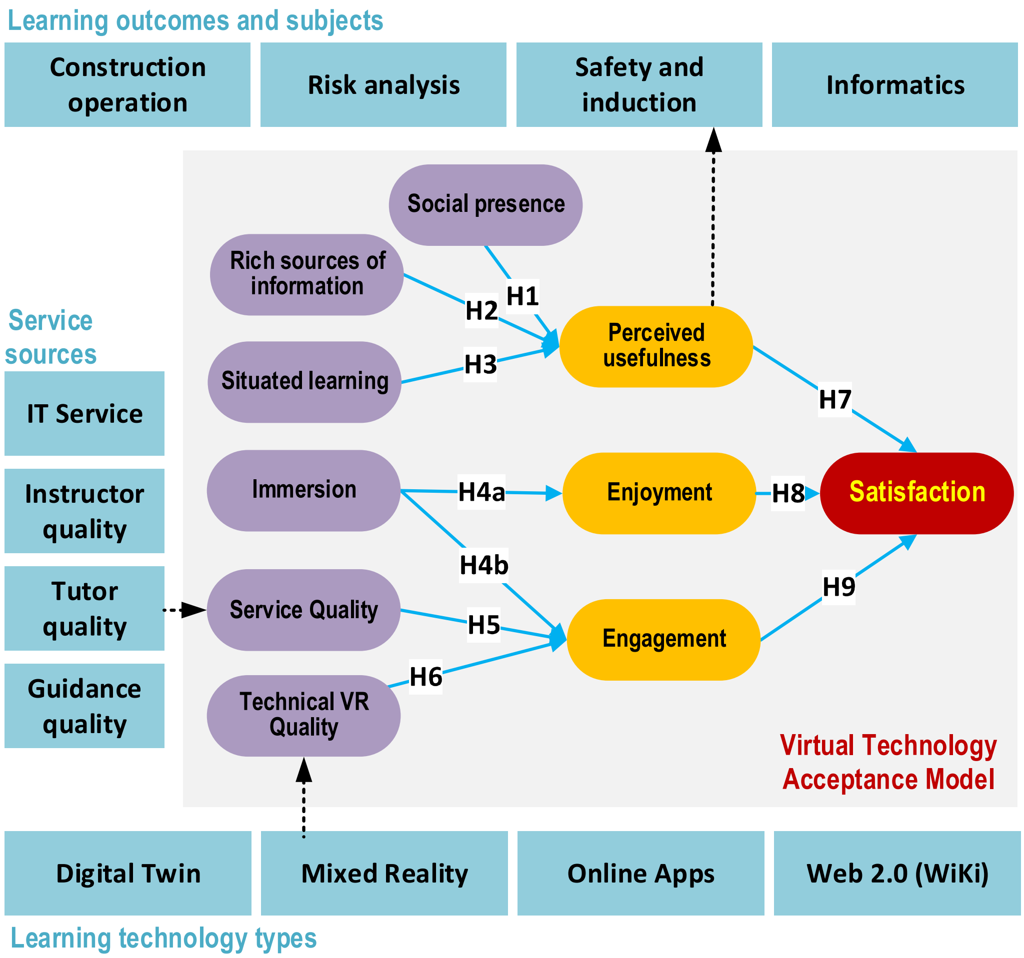


Figure 16. A proposed theoretical virtual technology acceptance model developed based on interviews, for measuring VR acceptance by participants.

Research question 2: The proposed theoretical model also can be modified for 360-degree video applications. For example, the following variables can be examined in modelling:

- Explore and comprehend construction site layout planning [76]
- Experience the site layout from a site manager perspective [76]
- Enhance the emotions felt by students
- Present a detailed scenario that can be consulted by students repeatedly if necessary [76]
- Identify risks and hazards related to different activities such as drilling rig. A survey can be conducted based on 'active processing assumption' that suggest meaningful learning may occur if students learners can organise the learning material into a coherent structure, and allow them to integrate it with their prior knowledge about the event [76,82]. The survey based on Slater-Usoh-Steed (SUS) [83,84] can be adopted for assessing the level of presence experience. At the same time, students use 360 videos and can rate their experience in terms of their feeling of "being there, realism, involvement" [85]. Another recommended model is Witmer and

Singer presence (WS) measuring student involvement, sensory fidelity, adaptation or immersion, and interface quality [86]. The post media questionnaire (PMQ) is also useful to design a question for assessing the emotions caused by 360 video exposure in terms of "anxiety, disgust, anger, fury, surprise, relax, happiness, and sadness" [85].

Research question 3: How technology-enhanced education and pedagogical design to increase students learning in practical courses, software teaching, and skilled-based subjects in construction. How different elements of virtual online learning will contribute to collaborative thinking and will enhance students' engagement in software teaching and skilled-based subjects. There is a wide range of new technologies such as light detection and ranging (Lidar) tools [69,81,87–90], data mining [91], deep/machine learning [92,93], geographic information systems (GIS) [70,94], different types of laser scanners [88,95–97], virtual reality, three-dimensional printing [98], building information modeling (BIM) [99–102] and digital twin [68,103] that should be practised and learned by students and new employees. The challenging questions concern 'how technology can be used to teach technology'. These tools are often required for employment. Recent education studies recommended that education mode and assessments should be reframed for enhancing students' skills for employability [104]. The questions raise how digital twin and new teaching tools can help educators to 'implement the constructive alignment model' in different subjects [105]. Some relevant research includes the work of Boje et al. [103] and da Motta Gaspar et al. [106] who intended to address this type of question. Still, both education technology and construction digital technologies are advancing, and thus training approaches are required to be modified based on empirical investigations over time.

Research question 4: Some studies tried to use virtual or tangible replicas as a part of digital twin in archaeology using manipulation interfaces (e.g., SketchFab) in archaeology [107]. However, there was no usage of the digital twin in ACE education. Future studies can continue the use of the digital twin in a different context, transferring different concepts to students how the digital twin and a virtual replica can contribute to education transformation and enhance the immersive environment and students' engagements. How can the learning management system (LMS), including Moodle or collaborative Blackboard, as well as communication tools such as Zoom or Microsoft Teams, support these digital simulations. Constructive alignment is challenging when using a digital device. Most of the technologies offer exciting features, but the question is if they will help the instructor to align further the course learning outcome, activities, and assessments. In fact, the question that should be addressed concerns on how virtual technology is useful in implementing the constructive alignment model in teaching.

6. Conclusions

This project was aimed to present a set of novel technologies and practices, which was used for developing a digital pedagogy. They introduced technologies that enable instructors to monitor students' performance and give them immediate feedback in large classes, where the instructor cannot trace students using traditional approaches. The outcomes and delivered materials discussed in this paper covered different topics such as the 'introduction to a construction project' (e.g., UNSW campus and the selected building) as the first section of the PAR, 'sequences of foundation construction' as the second part of the PAR, and 'design construction processes for tunnelling as a specific activity' covered by VTBM. One of the primary purposes of these online mixed reality modules is to improve the learning experience by allowing students to understand how a tunnel boring machine works, and to become familiar with the different structural elements of a building based on practical experience in a virtual environment. The technologies introduced in this paper offer an opportunity to engage students in acquiring and retaining their knowledge, learn practical knowledge of running an excavator or planning for the excavation process in a construction site as well as improving their teamwork and social skills.

The six primary online education tools were developed and presented in this paper by examining many scenario developments, group discussions, evaluations of initial versions, implementation, and revisions. There was considerable feedback from students to assess each module before the last updated version was finalised. The development team included experts with different backgrounds, lecturers, students, educational development specialists, industry practitioners, and technology experts. Lecturers, students, and educational development experts were involved in designing the modules to underline its usefulness to instructors and learners. Industry practitioners brought first-hand knowledge and case study information into the modules rather than the theoretical information available in textbooks. Details of each education module were discussed in the paper.

This paper is a step forward towards the implementation of a fully online immersive teaching experience in construction, while there are limited practices of the digital twin in the construction education context. The paper presents the potential of online mixed reality in construction education while it has not been thoroughly investigated in line with the advances made in technology development. The recent growth in virtual reality hardware and devices has increased the applicability of mixed reality practices and their ease of use.

The paper also presented the role-play group project model, including individuals and group activities and tasks. The model is implemented on an online platform. The application is beyond 'sharing information in WiKi', and it helps the student to assist each other, give immediate feedback, and correct each other on the report. The implementation of the model has many advantages comparing the current methods discussed, such as:

(1) The instructor can monitor groups or individuals whether they are continually making progress, and which group is waiting for the last week of the semester to do the job
(2) Students can get immediate feedback from their team members and the instructor (lecturer and tutors)
(3) The instructor can see the history of all changes, edits, and corrections have done by students.

The contributions of this paper are to extend the body of knowledge in building construction by presenting novel technological applications of digital twin and mixed reality in the construction context. The forms and design development process also present theoretical factors influencing the learning construction process, which increases the learner competency.

The implications of this paper are to present a novel technological approach for building and tunnelling construction education and professional training. Construction project managers can use this approach for two purposes: project induction and training construction operation to novice practitioners. This paper also clarifies the value for designing new online interactive tools for building construction, discusses the effectiveness of online tools in everyday teaching and learning, and raises awareness of online interactive applications in construction education and businesses. This paper presents the outcome of expensive experimentation, which is extremely valuable to construction educators who are from older generations to digital natives. This paper introduces practices to these educators who are accustomed to plan, work, and interact with others in a physical world rather than a virtual environment. However, the technologies presented in this paper can be counted as a subjective norm or expectations of the digital natives. They have been using online apps and virtual game-based environments from their early years. This paper also suggests a plan for future studies, including valuable research issues discussed in the discussion section. The article was limited to presenting potential solutions to the need for virtual reality modules, and the qualitative study limited to examine three modules. Thus, more empirical investigations are required, including surveys, to evaluate each module using a larger group of participants familiar with the virtual apps and tools presented in this paper.

Funding: This research received no external funding.

Acknowledgments: The practice supported by the Scientia Education Investment Fund (SEIF 2018-2019) and Research Infrastructure Scheme (RIS) at the University of New South Wales, Sydney. With thanks to the participation of instructors, students, technical practitioners, and industry partners and the Construction Mixed Reality Development (CONXR UNSW) stakeholders.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sepasgozar, S.M.E.; Bliemel, M.J.; Wang, C. A Flipped classroom model to teach skill-based contents for a large construction technology course. In Proceedings of the 40th AUBEA 2016, Radical Innovation in the Built Environment, Cairns, Australia, 6–8 July 2016.
2. Liu, J.; Wang, Q.; Liang, S.; Zhang, Z. Design of virtual reality combined with blended experimental teaching mode. In Proceedings of the 2019 International Conference on Advanced Education Research and Modern Teaching (AERMT 2019), Jinan, China, 28–29 September 2019.
3. Lee, V.W.; Hodgson, P.; Chan, C.-S.; Fong, A.; Cheung, S.W. Optimising the learning process with immersive virtual reality and non-immersive virtual reality in an educational environment. *Int. J. Mob. Learn. Organ.* **2020**, *14*, 21–35. [CrossRef]
4. Lee, A.L.; DeBest, M.; Koeniger-Donohue, R.; Strowman, S.R.; Mitchell, S.E. The feasibility and acceptability of using virtual world technology for interprofessional education in palliative care: A mixed methods study. *J. Interprof. Care* **2019**, 1–11. [CrossRef] [PubMed]
5. Bashabsheh, A.K.; Alzoubi, H.H.; Ali, M.Z. The application of virtual reality technology in architectural pedagogy for building constructions. *Alex. Eng. J.* **2019**, *58*, 713–723. [CrossRef]
6. Wang, R.; Lowe, R.; Newton, S.; Kocaturk, T. Task complexity and learning styles in situated virtual learning environments for construction higher education. *Autom. Constr.* **2020**, *113*, 103148. [CrossRef]
7. Eiris Pereira, R.; Gheisari, M. Site visit application in construction education: A descriptive study of faculty members. *Int. J. Constr. Educ. Res.* **2019**, *15*, 83–99. [CrossRef]
8. Gao, Y.; Gonzalez, V.A.; Yiu, T.W. The effectiveness of traditional tools and computer-aided technologies for health and safety training in the construction sector: A systematic review. *Comput. Educ.* **2019**, *138*, 101–115. [CrossRef]
9. Jensen, L.; Konradsen, F. A review of the use of virtual reality head-mounted displays in education and training. *Educ. Inf. Technol.* **2018**, *23*, 1515–1529. [CrossRef]
10. Feng, Z.; González, V.A.; Amor, R.; Lovreglio, R.; Cabrera-Guerrero, G. Immersive virtual reality serious games for evacuation training and research: A systematic literature review. *Comput. Educ.* **2018**, *127*, 252–266. [CrossRef]
11. Checa, D.; Bustillo, A. Advantages and limits of virtual reality in learning processes: Briviesca in the fifteenth century. *Virtual Real.* **2020**, *24*, 151–161. [CrossRef]
12. Han, S.; Capraro, R.; Capraro, M.M. How science, technology, engineering, and mathematics (STEM) project-based learning (PBL) affects high, middle, and low achievers differently: The impact of student factors on achievement. *Int. J. Sci. Math. Educ.* **2015**, *13*, 1089–1113. [CrossRef]
13. Basilotta Gómez-Pablos, V.; Martín del Pozo, M.; García-Valcárcel Muñoz-Repiso, A. Project-based learning (PBL) through the incorporation of digital technologies: An evaluation based on the experience of serving teachers. *Comput. Hum. Behav.* **2017**, *68*, 501–512. [CrossRef]
14. Johnson, D.W.; Johnson, R.T. An educational psychology success story: Social interdependence theory and cooperative learning. *Educ. Res.* **2009**, *38*, 365–379. [CrossRef]
15. James, K.; Humphrey, G.; Vilis, T.; Corrie, B.; Baddour, R.; Goodale, M. "Active" and "passive" learning of three-dimensional object structure within an immersive virtual reality environment. *Behav. Res. Methods Instrum. Comput.* **2002**, *34*, 383–390. [CrossRef] [PubMed]
16. Horne, M.; Thompson, E.M. The role of virtual reality in built environment education. *J. Educ. Built Environ.* **2008**, *3*, 5–24. [CrossRef]
17. Wang, P.; Wu, P.; Wang, J.; Chi, H.-L.; Wang, X. A critical review of the use of virtual reality in construction engineering education and training. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1204. [CrossRef]
18. Klotz, L.; Grant, D. A balanced view of sustainability in civil engineering and construction. In Proceedings of the Construction Research Congress, Seattle, WA, USA, 5–7 April 2009; pp. 1338–1347.
19. Lin, C.-C.; Tsai, C.-C. The relationships between students' conceptions of learning engineering and their preferences for classroom and laboratory learning environments. *J. Eng. Educ.* **2009**, *98*, 193. [CrossRef]
20. Yeary, M.; Yu, T.-Y.; Palmer, R.; Biggerstaff, M.; Fink, L.; Ahern, C.; Tarp, K. A hands-on, interdisciplinary laboratory program and educational model to strengthen a radar curriculum for broad distribution. *J. Adv. Eng. Educ* **2007**, *1*, 1–23.

21. Leicht, R.M.; Lewis, A.; Riley, D.R.; Messner, J.I.; Darnell, B. Assessing traits for success in individual and team performance in an engineering course. In Proceedings of the ASCE Construction Research Congress, Seattle, WA, USA, 5–7 April 2009; pp. 5–7.
22. Toto, R.; Nguyen, H. Flipping the work design in an industrial engineering course. In Proceedings of the Frontiers in Education Conference, FIE'09. 39th IEEE, San Antonio, TX, USA, 18–21 October 2009; pp. 1–4.
23. Monson, M.C.; Homayouni, H.; Dossick, C.; Anderson, A. Improving the understanding of BIM concepts through a flipped learning lab environment: A work in progress. In Proceedings of the ASEE 122nd Annual Conference & Exposition, Seattle, WA, USA, 14–17 June; pp. 14–17.
24. Wang, S.; Cheah, C.Y.; Chew, D.A. Dynamics of strategic management in the Chinese construction industry. *Manag. Decis.* **2005**. [CrossRef]
25. McWhirter, N.; Shealy, T. Case-based flipped classroom approach to teach sustainable infrastructure and decision-making. *Int. J. Constr. Educ. Res.* **2020**, *16*, 3–23. [CrossRef]
26. Lee, S.J.; Ngampornchai, A.; Trail-Constant, T.; Abril, A.; Srinivasan, S. Does a case-based online group project increase students' satisfaction with interaction in online courses? *Act. Learn. High. Educ.* **2016**, *17*, 249–260. [CrossRef]
27. Finlay, S.-J.; Faulkner, G. Tête à tête Reading groups and peer learning. *Act. Learn. High. Educ.* **2005**, *6*, 32–45. [CrossRef]
28. Bahar-Özvaris̨, S.e.; Çetin, F.Ç.; Turan, S.; Peters, A.S. Cooperative learning: A new application of problem-based learning in mental health training. *Med Teach.* **2006**, *28*, 553–557. [CrossRef] [PubMed]
29. Herrmann, K.J. The impact of cooperative learning on student engagement: Results from an intervention. *Act. Learn. High. Educ.* **2013**, *14*, 175–187. [CrossRef]
30. Hall, D.; Buzwell, S. The problem of free-riding in group projects: Looking beyond social loafing as reason for non-contribution. *Act. Learn. High. Educ.* **2012**, *14*, 37–49. [CrossRef]
31. Hassanien, A. A qualitative student evaluation of group learning in higher education. *High. Educ. Eur.* **2007**, *32*, 135–150. [CrossRef]
32. Radianti, J.; Majchrzak, T.A.; Fromm, J.; Wohlgenannt, I. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Comput. Educ.* **2020**, *147*, 103778. [CrossRef]
33. Lee, E.A.-L.; Wong, K.W.; Fung, C.C. How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Comput. Educ.* **2010**, *55*, 1424–1442.
34. Davis, F.D. User acceptance of information technology; system characteristics, user perceptions and behavioral impacts. *Int. J. Manag. Mach. Stud.* **1993**, 475–487. [CrossRef]
35. Sepasgozar, S.M.E.; Bernold, L.E. A Technology pre-adoption model for construction. In Proceedings of the 37th Annual Conference of Australasian University Building Educators Association (AUBEA), Sydney, Australia, 30 September–3 October 2008.
36. McNamara, A.J.; Sepasgozar, S.M. Developing a theoretical framework for intelligent contract acceptance. *Constr. Innov.* **2020**. [CrossRef]
37. Garone, A.; Pynoo, B.; Tondeur, J.; Cocquyt, C.; Vanslambrouck, S.; Bruggeman, B.; Struyven, K. Clustering university teaching staff through UTAUT: Implications for the acceptance of a new learning management system. *Br. J. Educ. Technol.* **2019**, *50*, 2466–2483. [CrossRef]
38. Revythi, A.; Tselios, N. Extension of Technology Acceptance Model by using System Usability Scale to assess behavioral intention to use e-learning. *Educ. Inf. Technol.* **2019**, *24*, 2341–2355. [CrossRef]
39. Nakarada-Kordic, I.; Reay, S.; Bennett, G.; Kruse, J.; Lydon, A.-M.; Sim, J. Can virtual reality simulation prepare patients for an MRI experience? *Radiography* **2019**. [CrossRef]
40. Kim, H.K.; Park, J.; Choi, Y.; Choe, M. Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. *Appl. Ergon.* **2018**, *69*, 66–73. [CrossRef]
41. Bergmann, J.; Sams, A. *Flipped learning: Gateway to Student Engagement*; International Society for Technology in Education: Washington, DC, USA, 2014.
42. Balan, P.; Clark, M.; Restall, G. Preparing students for Flipped or Team-Based Learning methods. *Educ. Train.* **2015**, *57*, 639–657. [CrossRef]
43. Bliemel, M.J. Lessons learned from an inside-out flip in entrepreneurship education. *Small Enterp. Res.* **2014**, *21*, 117–128. [CrossRef]

44. Goldfinch, J.; Raeside, R. Development of a peer assessment technique for obtaining individual marks on a group project. *Assess. Eval. High. Educ.* **1990**, *15*, 210–231. [CrossRef]
45. Lejk, M.; Wyvill, M. Peer assessment of contributions to a group project: A comparison of holistic and category-based approaches. *Assess. Eval. High. Educ.* **2001**, *26*, 61–72. [CrossRef]
46. Brooks, C.M.; Ammons, J.L. Free riding in group projects and the effects of timing, frequency, and specificity of criteria in peer assessments. *J. Educ. Bus.* **2003**, *78*, 268–272. [CrossRef]
47. Juhaňák, L.; Zounek, J.; Rohlíková, L. Using process mining to analyze students' quiz-taking behavior patterns in a learning management system. *Comput. Hum. Behav.* **2019**, *92*, 496–506. [CrossRef]
48. Biasutti, M.; EL-Deghaidy, H. Interdisciplinary project-based learning: An online wiki experience in teacher education. *Technol. Pedagog. Educ.* **2015**, *24*, 339–355. [CrossRef]
49. Elmahadi, I.; Osman, I. A study of the Sudanese students' use of collaborative tools within moodle learning management system. In Proceedings of the IST-Africa Conference and Exhibition (IST-Africa), Nairobi, Kenya, 29–31 May 2013; pp. 1–8.
50. Sonego, A.H.S.; do Amaral, É.M.H.; Nunes, F.B.; Voss, G.B. Use of Moodle as a tool for collaborative learning: A study focused on wiki. *IEEE Rev. Iberoam. Tecnol. Aprendiz.* **2014**, *9*, 17–21. [CrossRef]
51. Gutiérrez-Braojos, C.; Montejo-Gamez, J.; Marin-Jimenez, A.; Campaña, J. Hybrid learning environment: Collaborative or competitive learning? *Virtual Real.* **2019**, *23*, 411–423. [CrossRef]
52. Cooper, G.; Park, H.; Nasr, Z.; Thong, L.; Johnson, R. Using virtual reality in the classroom: Preservice teachers' perceptions of its use as a teaching and learning tool. *Educ. Media Int.* **2019**, *56*, 1–13. [CrossRef]
53. Monahan, T.; McArdle, G.; Bertolotto, M. Virtual reality for collaborative e-learning. *Comput. Educ.* **2008**, *50*, 1339–1353. [CrossRef]
54. Zheng, L.; Zhang, X.; Gyasi, J.F. A literature review of features and trends of technology-supported collaborative learning in informal learning settings from 2007 to 2018. *J. Comput. Educ.* **2019**, *6*, 529–561. [CrossRef]
55. Dominic, M.; Francis, S.; Pilomenraj, A. E-learning in web 3.0. *Int. J. Mod. Educ. Comput. Sci.* **2014**, *6*, 8. [CrossRef]
56. Männistö, M.; Mikkonen, K.; Kuivila, H.M.; Virtanen, M.; Kyngäs, H.; Kääriäinen, M. Digital collaborative learning in nursing education: A systematic review. *Scand. J. Caring Sci.* **2019**. [CrossRef] [PubMed]
57. Halavais, A. Computer-Supported Collaborative Learning. *Int. Encycl. Commun. Theory Philos.* **2016**, 1–5. [CrossRef]
58. Blau, I.; Shamir-Inbal, T.; Avdiel, O. How does the pedagogical design of a technology-enhanced collaborative academic course promote digital literacies, self-regulation, and perceived learning of students? *Internet High. Educ.* **2020**, *45*, 100722. [CrossRef]
59. Sepasgozar, S.M.; Davis, S.; Loosemore, M.; Bernold, L. An investigation of modern building equipment technology adoption in the Australian construction industry. *Eng. Constr. Archit. Manag.* **2018**. [CrossRef]
60. Sepasgozar, S.M.; Davis, S.R.; Li, H.; Luo, X. Modeling the Implementation Process for New Construction Technologies: Thematic Analysis Based on Australian and US Practices. *J. Manag. Eng.* **2018**, *34*, 05018005. [CrossRef]
61. Sepasgozar, S.M.; Davis, S. Construction Technology Adoption Cube: An Investigation on Process, Factors, Barriers, Drivers and Decision Makers Using NVivo and AHP Analysis. *Buildings* **2018**, *8*, 74. [CrossRef]
62. Mak, T.M.; Iris, K.; Wang, L.; Hsu, S.-C.; Tsang, D.C.; Li, C.; Yeung, T.L.; Zhang, R.; Poon, C.S. Extended theory of planned behaviour for promoting construction waste recycling in Hong Kong. *Waste Manag.* **2019**, *83*, 161–170. [CrossRef]
63. Horton, J.; Macve, R.; Struyven, G. Qualitative research: Experiences in using semi-structured interviews. In *The Real Life Guide to Accounting Research*; Elsevier: Amsterdam, The Netherlands, 2004; pp. 339–357.
64. Sepasgozar, S.M.E.; Loosemore, M.; Davis, S.R. Conceptualising information and equipment technology adoption in construction A critical review of existing research. *Eng. Constr. Archit. Manag.* **2016**, *23*, 158–176. [CrossRef]
65. Rogers, E.M. *Diffusion of Innovations*, 4th ed.; Free Press: New York, NY, USA, 1995.
66. Sepasgozar, S. Footing Construction Process [Immersive Environment for Teaching]. YouTube: Sydney. 2020. Available online: https://www.youtube.com/watch?v=aZFGxmVAbFM (accessed on 1 July 2020).
67. Sepasgozar, S. Site Layout [Immersive Teaching] Sydney. 2020. Available online: https://www.youtube.com/watch?v=5qJnLjjr2jY (accessed on 1 July 2020).

68. Shirowzhan, S.; Tan, W.; Sepasgozar, S.M. *Digital Twin and CyberGIS for Improving Connectivity and Measuring the Impact of Infrastructure Construction Planning in Smart Cities*; Multidisciplinary Digital Publishing Institute: Basel, Switzerland, 2020.
69. Shirowzhan, S.; Sepasgozar, S.M.E.; Li, H.; Trinder, J. Spatial compactness metrics and Constrained Voxel Automata development for analyzing 3D densification and applying to point clouds: A synthetic review. *Autom. Constr.* **2018**, *96*, 236–249. [CrossRef]
70. Shirowzhan, S.; Sepasgozar, S.M. Spatial analysis using temporal point clouds in advanced GIS: Methods for ground elevation extraction in slant areas and building classifications. *ISPRS Int. J. Geo Inf.* **2019**, *8*, 120. [CrossRef]
71. Venkatesh, V.; Morris, M.G.; Davis, G.B.; Davis, F.D. User acceptance of information technology: Toward a unified view. *MIS Q. Manag. Inf. Syst.* **2003**, *27*, 425–478. [CrossRef]
72. Lee, J.; Kim, J.; Ahn, J.; Woo, W. Context-aware risk management for architectural heritage using historic building information modeling and virtual reality. *J. Cult. Herit.* **2019**, *38*, 242–252. [CrossRef]
73. Shi, Y.; Du, J.; Ahn, C.R.; Ragan, E. Impact assessment of reinforced learning methods on construction workers' fall risk behavior using virtual reality. *Autom. Constr.* **2019**, *104*, 197–214. [CrossRef]
74. Jantakoon, T.; Wannapiroon, P.; Nilsook, P. Virtual Immersive Learning Environments (VILEs) Based on Digital Storytelling to Enhance Deeper Learning for Undergraduate Students. *High. Educ. Stud.* **2019**, *9*, 144–150. [CrossRef]
75. Templeton, C.; Kessinger, M.W. Virtual Reality: Learning by Seeing in 3D. In *Handbook of Research on Software for Gifted and Talented School Activities in K-12 Classrooms*; IGI Global: Hershey, PA, USA, 2020; pp. 94–119.
76. Violante, M.G.; Vezzetti, E.; Piazzolla, P. Interactive virtual technologies in engineering education: Why not 360° videos? *Int. J. Interact. Des. Manuf. (Ijidem)* **2019**, *13*, 729–742. [CrossRef]
77. Raghavan, R.; Rao, P. *Accenture Extended Reality (XR): Immersive Learning for the Future Workforce*; IEEE: New York, NY, USA, 2018.
78. Scherer, R.; Siddiq, F.; Tondeur, J. The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Comput. Educ.* **2019**, *128*, 13–35. [CrossRef]
79. Venkatesh, V. Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model. *Inf. Syst. Res.* **2000**, *11*, 342–365. [CrossRef]
80. Sepasgozar, S.M.; Hawken, S.; Sargolzaei, S.; Foroozanfa, M. Implementing citizen centric technology in developing smart cities: A model for predicting the acceptance of urban technologies. *Technol. Forecast. Soc. Chang.* **2019**, *142*, 105–116. [CrossRef]
81. Sepasgozar, S.; Shirowzhan, S.; Wang, C.C. A Scanner Technology Acceptance Model for Construction Projects. *Procedia Eng.* **2017**, 1237–1246. [CrossRef]
82. Wittrock, M.C. Generative Processes of Comprehension. *Educ. Psychol.* **1989**, *24*, 345–376. [CrossRef]
83. Usoh, M.; Catena, E.; Arman, S.; Slater, M. Using presence questionnaires in reality. *Presence Teleoper. Virtual Environ.* **2000**, *9*, 497–503. [CrossRef]
84. Youngblut, C.; Huie, O. The relationship between presence and performance in virtual environments: Results of a VERTS study. In Proceedings of the IEEE Virtual Reality, Los Angeles, CA, USA, 22–26 March 2003; pp. 277–278.
85. Pallavicini, F.; Cipresso, P.; Raspelli, S.; Grassi, A.; Serino, S.; Vigna, C.; Triberti, S.; Villamira, M.; Gaggioli, A.; Riva, G. Is virtual reality always an effective stressors for exposure treatments? Some insights from a controlled trial. *BMC Psychiatry* **2013**, *13*, 52. [CrossRef]
86. Witmer, B.G.; Jerome, C.J.; Singer, M.J. The factor structure of the presence questionnaire. *Presence Teleoper. Virtual Environ.* **2005**, *14*, 298–312. [CrossRef]
87. Sepasgozar, S.M.; Wang, C.; Shirowzhan, S. Challenges and Opportunities for Implementation of Laser Scanners in Building Construction. In Proceedings of the 33rd International Symposium on Automation and Robotics in Construction (ISARC 2016), Auburn, Alabama, USA, 18 July 2016; pp. 742–751.
88. Shirowzhan, S.; Sepasgozar, S.; Liu, C. Monitoring physical progress of indoor buildings using mobile and terrestrial point clouds. In Proceedings of the Construction Research Congress, New Orleans, Louisiana, USA, 16 May 2018.
89. Shirowzhan, S.; Lim, S.; Trinder, J. Enhanced autocorrelation-based algorithms for filtering airborne lidar data over urban areas. *J. Surv. Eng.* **2016**, *142*, 04015008. [CrossRef]

90. Shirowzhan, S.; Trinder, J. Building classification from lidar data for spatio-temporal assessment of 3D urban developments. *Procedia Eng.* **2017**, *180*, 1453–1461. [CrossRef]
91. Shirowzhan, S.; Lim, S.; Trinder, J.; Li, H.; Sepasgozar, S.M.E. Data mining for recognition of spatial distribution patterns of building heights using airborne lidar data. *Adv. Eng. Inform.* **2020**, *43*, 101033. [CrossRef]
92. Shirowzhan, S.; Sepasgozar, S.M.; Li, H.; Trinder, J.; Tang, P. Comparative analysis of machine learning and point-based algorithms for detecting 3D changes in buildings over time using bi-temporal lidar data. *Autom. Constr.* **2019**, *105*, 102841. [CrossRef]
93. Zhong, B.; Xing, X.; Luo, H.; Zhou, Q.; Li, H.; Rose, T.; Fang, W. Deep learning-based extraction of construction procedural constraints from construction regulations. *Adv. Eng. Inform.* **2020**, *43*, 101003. [CrossRef]
94. Shirowzhan, S.; Sepasgozar, S.M.E.; Zaini, I.; Wang, C. An integrated GIS and Wi-Fi based locating system for improving construction labor communications. In Proceedings of the 34th International Symposium on Automation and Robotics in Construction (ISARC 2017), Taipei, Taiwan, 1 January 2017; pp. 1052–1059.
95. Sepasgozar, S.M.; Forsythe, P.; Shirowzhan, S.; Norzahari, F. Scanners And photography: A combined framework. In Proceedings of the 40th Australasian Universities Building Education Association (AUBEA) 2016 Conference, Cairns, Australia, 8 July 2016; pp. 819–828.
96. Sepasgozar, S.M.; Forsythe, P.; Shirowzhan, S. Evaluation of terrestrial and mobile scanner technologies for part-built information modeling. *J. Constr. Eng. Manag.* **2018**, *144*, 04018110. [CrossRef]
97. Sepasgozar, S.; Lim, S.; Shirowzhan, S.; Kim, Y.; Nadoushani, Z.M. Utilisation of a New Terrestrial Scanner for Reconstruction of As-built Models: A Comparative Study. In Proceedings of the International Symposium on Automation and Robotics in Construction, Oulu, Finland, 15–18 June 2015.
98. Tahmasebinia, F.; Niemelä, M.; Ebrahimzadeh Sepasgozar, S.; Lai, T.; Su, W.; Reddy, K.; Shirowzhan, S.; Sepasgozar, S.; Marroquin, F. Three-Dimensional Printing Using Recycled High-Density Polyethylene: Technological Challenges and Future Directions for Construction. *Buildings* **2018**, *8*, 165. [CrossRef]
99. Shirowzhan, S.; Sepasgozar, S.M.E.; Edwards, D.J.; Li, H.; Wang, C. BIM compatibility and its differentiation with interoperability challenges as an innovation factor. *Autom. Constr.* **2020**, *112*, 103086. [CrossRef]
100. Zhao, P.A.; Wang, C.C. A comparison of using traditional cost estimating software and BIM for construction cost control. In Proceedings of the ICCREM 2014: Smart Construction and Management in the Context of New Technology, Kunming, China, 27–28 September 2014; pp. 256–264.
101. Sami Ur Rehman, M.; Thaheem, M.J.; Nasir, A.R.; Khan, K.I.A. Project schedule risk management through building information modelling. *Int. J. Constr. Manag.* **2020**, 1–11. [CrossRef]
102. Hosseini, M.R.; Chileshe, N.; Zuo, J.; Baroudi, B. Adopting global virtual engineering teams in AEC Projects. *Constr. Innov.* **2015**. [CrossRef]
103. Boje, C.; Guerriero, A.; Kubicki, S.; Rezgui, Y. Towards a semantic Construction Digital Twin: Directions for future research. *Autom. Constr.* **2020**, *114*, 103179. [CrossRef]
104. Ruge, G.; McCormack, C. Building and construction students' skills development for employability–reframing assessment for learning in discipline-specific contexts. *Archit. Eng. Des. Manag.* **2017**, *13*, 365–383. [CrossRef]
105. Ruge, G.; Tokede, O.; Tivendale, L. Implementing constructive alignment in higher education—Cross -institutional perspectives from Australia. *High. Educ. Res. Dev.* **2019**, *38*, 833–848. [CrossRef]
106. Da Motta Gaspar, J.A.; Ruschel, R.C.; Monteiro, E.Z. Integrated and collaborative architectural design: 10 years of experience teaching BIM. In *Advances in Informatics and Computing in Civil and Construction Engineering*; Springer: Berlin, Germany, 2019; pp. 865–872.
107. Pollalis, C.; Minor, E.J.; Westendorf, L.; Fahnbulleh, W.; Virgilio, I.; Kun, A.L.; Shaer, O. Evaluating learning with tangible and virtual representations of archaeological artifacts. In Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction, Stockholm, Sweden, 18–21 March 2018; pp. 626–637.

applied sciences

MDPI

Review

Digital Twin and Internet of Things—Current Standards Landscape

Michael Jacoby * and **Thomas Usländer**

Fraunhofer IOSB, Fraunhofer Institute of Optronics, System Technologies and Image Exploitation, 76131 Karlsruhe, Germany; thomas.uslaender@iosb.fraunhofer.de

* Correspondence: michael.jacoby@iosb.fraunhofer.de; Tel.: +49-721-6091-470

Received: 20 August 2020; Accepted: 15 September 2020 ; Published: 18 September 2020

Abstract: Industry 4.0 is revolutionizing industrial production by bridging the physical and the virtual worlds and further improving digitalization. Two essential building blocks in industry 4.0 are digital twins (DT) and the internet of things (IoT). While IoT is about connecting resources and collecting data about the physical world, DTs are the virtual representations of resources organizing and managing information and being tightly integrated with artificial intelligence, machine learning and cognitive services to further optimize and automate production. The concepts of DTs and IoT are overlapping when it comes to describing, discovering and accessing resources. Currently, there are multiple DT and IoT standards covering these overlapping aspects created by different organizations with different backgrounds and perspectives. With regard to interoperability, which is presumably the most important aspect of industry 4.0, this barrier needs to be overcome by consolidation of standards. The objective of this paper is to investigate current DT and IoT standards and provide insights to stimulate this consolidation. Overlapping aspects are identified and a classification scheme is created and applied to the standards. The results are compared, aspects with high similarity or divergence are identified and a proposal for stimulating consolidation is presented. Consensus between standards are found regarding the elements a resource should consist of and which serialization format(s) and network protocols to use. Controversial topics include which query language to use for discovery as well as if geo-spatial, temporal and historical data should be explicitly supported.

Keywords: digital twin; internet of things; interoperability; standardization

1. Introduction

The fourth industrial revolution and the underlying digital transformation, known as Industry 4.0, is revolutionizing industrial production. Following the third industrial revolution, also known as the Digital Revolution, Industry 4.0 is about optimization and automation of the previously introduced computerization by intelligent networking of machines and processes [1,2]. A central element is the integration of already-existing concepts such as the internet of things (IoT), digital twins (DT), artificial intelligence (AI), machine learning (ML), big data, and cognitive services and applying them to industrial production. Three of the integral design principles of Industry 4.0 are interconnection, information transparency, and decentralized decisions [3]. Interconnection refers to the ability of machines, devices, sensors, and people to connect and communicate with each other via the internet, which is the cornerstone of the IoT. Information transparency refers to the ability to collect, manage, and organize the vast amount of data from connected machines, devices, and sensors. This requirement is addressed by the concept of DT, which is 'a formal digital representation of some asset, process or system that captures attributes and behaviors of that entity suitable for communication, storage, interpretation or processing within a certain context' [4]. The requirement for decentralized decisions can be fulfilled by incorporating aspects of AI and ML into DTs to provide cognitive services and thereby enable better decision-making and increasing the level of automation as well as overall productivity. DTs with such enhanced capabilities are also referred to as cognitive twins (CT) [5].

The importance of industry 4.0 and its related concepts and technologies has been widely acknowledged [6–9]. Between 2012 and 2016, IoT (or the similar 'internet of everything') was considered one of the ten most important technology trends of the year by Gartner, followed by DT as one of the five most important technology trends between 2017 and 2019 [10–17]. In 2020, the most important technology trend according to Gartner is hyperautomation, which refers to the totality of automation efforts across an entire organization and is tightly coupled to industry 4.0 and DT as 'hyperautomation often results in the creation of a digital twin of the organization' [18].

It seems obvious that 'the rise of digital twins coincides with the rise of the IoT' [19] as only the vast amount of sensor data and device metadata from the IoT creates the requirement for organizing and managing all that information in an adequate way, which is realized by the concept of DTs. Although the relation between DT and IoT seems to be clear at first glance—IoT connects devices to the internet and collects data while the concept of DTs is used to structure and manage that data to be further used for optimization and automation with the help of AI and ML algorithms—they have some overlap and therefore are not clearly separable in some aspects.

A fundamental aspect of both IoT and DT is the fact that they center on resources and require a formal and machine-readable representation of these resources. Resources can be devices, sensors, actuators, or machines in the context of IoT or assets, processes, or systems in the context of DT. Figure 1 provides a visual representation of further resource-related functionality that is part of both domains. Besides a formal and machine-readable resource description created by a provider, they share the functionality for potential consumers to discover such resource descriptions and access the resources.

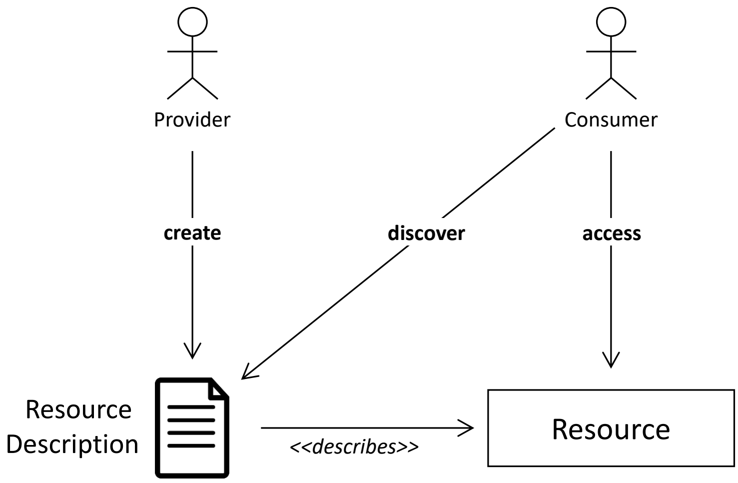


Figure 1. Functional overlap between internet of things and digital twins that is subject to standardization in both domains.

As industry 4.0 is about optimization on a large scale, interoperability not only within a factory or organization but along the whole value chain is essential and should be addressed globally [20,21]. Achieving interoperability on a global level requires formal standardization. Multiple SDOs are currently working on standardization of IoT or DT, e.g., the European Telecommunications Standards Institute (ETSI), the International Electrotechnical Commission (IEC), the Internet Engineering Task Force (IETF), the Industrial Internet Consortium (IIC), the International Organization for Standardization (ISO), the Open Geospatial Consortium, the Plattform Industrie 4.0, the World Wide Web Consortium (W3C), and many more. Many of the standards related to DT or IoT address issues that are unique to the domain, i.e., not part of the overlap between IoT and DT, for example, network protocols for wireless connectivity, which belongs to IoT but is not considered a relevant subject for standardization in the DT domain. Furthermore, different SDOs have different backgrounds and focuses, e.g., ETSI is rooted in the telecommunication domain, OGC focuses on geospatial information management, and the IIC is focused on accelerating the development and adoption of IoT technology in the industry. Several standards that focus on the three basic aspects common to both IoT and DT, resource description, resource discovery, and resource access, are either already existing or currently being developed by different SDOs in parallel.

Having multiple competing standards solving the same problem in different ways without aligning with each other massively hinders interoperability. As interoperability is the key requirement of industry 4.0 [21,22], this is an actual threat for the further development and success of industry 4.0. Previous work focuses on either the IoT or the DT domain or addresses the IoT only as a source of data for DTs, ignoring the fact that the domains actually overlap in functionality. A recent survey from Minerva et al. acknowledges the fact that interoperability plays an important role in industry 4.0 on IoT on the DT level and that 'interoperability and standardization will occur as part of the evolution' [23]. Although the relationship and alignment between existing and upcoming IoT and DT standards have been discussed in multiple SDO groups, to the best of the author's knowledge, there has not been any published formalized and structured analysis and comparison. The objective of this paper is to narrow this gap and stimulate the evolution of standards by identifying how well current standards from both the IoT and the DT domain are aligned regarding resource description, discovery and access. This not only provides an overview of the current state of the art on DT interoperability, but can also nourish further consolidation of standards to improve interoperability in industry 4.0.

Market research report predicts that until 2027, using DTs will become standard for IoT applications and IoT platforms will support the creation of DTs. At the same time, the total IoT market is predicted to rise from $465 million in 2019 to $1.5 trillion in 2030 whereby 40% thereof depend on successfully addressing interoperability [24–26]. As 'Industry 4.0 is only possible with the digital twin' [27], identifying and resolving ambiguities in and closing the gap between standards and thereby fostering interoperability is of relevance for the success of industry 4.0.

The reminder of the paper is organized as follows. The classification scheme for IoT and DT standards is presented in Section 2. In Section 3, multiple standards from the domains IoT and DT are presented and analyzed based on the previously introduce classification scheme. This is followed by a summary including a comparison matrix discussing the findings and conclusions in Sections 4 and 5, respectively.

2. Classification Schema

Both IoT and DT center on resources. In the context of IoT, resources are internet-connected devices, whereby communication between consumer and device can happen either directly or via some kind of software system. In the context of DT, resources are defined in a broader sense, e.g., as assets, devices, and physical or virtual entities. Both concepts share the vision that communication between resources should happen mostly without human interaction, i.e., machine-to-machine (M2M).

Figure 1 depicts basic functions that are both required for realizing IoT and DT. As M2M communication requires a formal representation of the resource to interact with, the first step that needs to happen in such a scenario is always the creation of a resource description. A resource description is exactly that formal representation required and describes a resource with all its capabilities and interfaces that should be exposed to others. The creator of a resource description has the role of the provider. Once a resource description is created, it can be discovered by a consumer. A consumer can be a software system (i.e., M2M) or a human. Different modes of resource discovery can be supported, e.g., peer-to-peer (P2P) or directory-style (see Section 2.2). Once discovered, a consumer can parse/read the resource description and extract the information needed to access the described resource.

Although these are basic functions used in many other domains, these are the very functions enabling the IoT and DT to revolutionize many aspects of today's IT landscape.

In the following, we present a classification schema for IoT and DT standards centered on the identified three main functions: resource description, resource discovery, and resource access.

2.1. Resource Description

Having a formal and machine-readable resource description is essential for IoT and DT, as it enables consumers to gain knowledge about available resources and defines how to interact with them. Resource descriptions are typically defined using a specific language or (meta) model, which can differ in many aspects, e.g., complexity, serialization formats, etc. In this paper, we discuss the following different classification criteria in the context of resource description.

2.1.1. Resource Term

What term is the central element of the language/(meta) model? Strictly speaking, this is not a real classification criterion but a helpful fact to find your way around the different terminologies used in different standards and platforms.

2.1.2. Model Type

Resource descriptions are expressed in the terms of a model (sometimes also referred to as a language). Depending on the expressiveness and generalizability, this model can be categorized as a different level following the Object Management Group (OMG) basic concept of multilevel metamodeling [28]. A detailed meta model hierarchy for IoT and DT is present in Figure 2. For each level, its content is described in the terms of the level above, whereas the top-level (*M3*) is typically self-describing. For example, a grounding in *M3* could be RDF(S) (Resource Description Format Schema, optionally in combination with RDF Schema), M2 could be the Web Ontology Language (OWL), M1 could contain multiple (cross-)domain or application models like the Semantic Sensor Network (SSN) Ontology and M0 would contain the actual application data. M1 and M2 are often a hierarchy or network themselves built by re-using existing (and typically standardized) models combined with custom extension In Figure 2, this is shown for M1 in an exemplary manner to reflect the hierarchical organization of models typically used in IoT and DT.

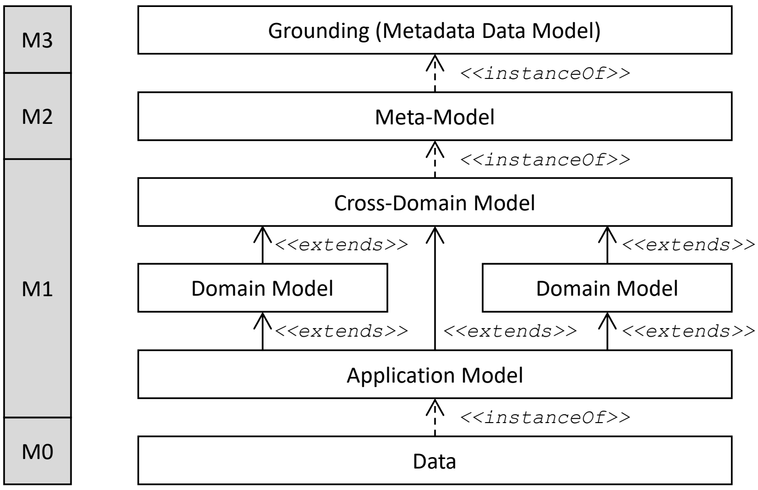


Figure 2. The metamodel hierarchy based on the Object Management Group basic idea of multilevel metamodeling [28] with an expanded M1 layer to reflect hierarchical organization of models typical for the internet of things and digital twins.

2.1.3. Resource Identification

Unique identification of any resource across system boundaries is essential in such highly distributed environments as IoT and DT (DTs may be passed along the supply chain along with (intermediate) products). Creating and keeping track of globally unique identifiers is not a trivial task and itself subject to standardization. Supporting different types of resource identifiers is also important when adding semantics through referencing external resources. Internationalized Resource

Identifiers (IRIs) are essential when working with ontologies as it is one of the basic building blocks of the Semantic Web. Other important external sources for semantics are international standards, e.g., eCl@ss [29], which often use International Registration Data Identifiers (IRDIs) to identify terms or concepts.

2.1.4. Type System

None of the considered standards provides a 'ready-to-use' application model but rather offer one or more models of a higher level (see Figure 2). This means that for a real-world application, a provider must define an application model based on the given higher-level models defining the concrete resource classes and their properties. This requires a type system defining the allowed/supported data types. Most standards re-use existing type systems, e.g., from XSD, JSON, or RDF, either directly or with slight restrictions and/or extensions.

2.1.5. Serialization Formats

To be exchanged between provider and consumer, a resource description needs to be serialized. This can be done in different formats like XML, JSON, JSON-LD, RDF or many others. Theoretically, most of these formats are interchangeable which also manifests in the fact that some standards support multiple serialization formats. However, the choice of serialization format can still influence other aspects, because e.g., a format might be coupled to a communication protocol or the type systems of the serialization formats might be difficult to match.

2.1.6. Resource Elements

The structure and expressiveness of a resource description language are one of the most important aspects of this classification. It defines what kind of information about a resource description can contain. The purpose of a resource description is to provide a consumer with all the necessary information to understand what this resource is and how to interact with it. It can be divided into two parts, one describing the capabilities of a resource and another one describing how to access it. The description of capabilities is very similar to the concept of interfaces in high-level programming languages and comes down to three types of elements: properties, functions, and events.

Properties refer to attributes of the resource and allow reading/writing or subscribing to them. Functions allow executing some logic on the resource supporting optional input and output parameters. In the context of IoT and DT, the term function is rarely used but instead the same concept is referred to as 'service'. Events are notifications about the occurrence of a state (change) or an action, e.g., 'battery level is critical' or 'button X has been pressed'.

2.1.7. Kinds of Data

The great variety of different use cases and application domains of IoT and DT require coping with different kind of data. The Big Data Value Association (BDVA) identified six different types of data, including time series, geo-spatial, temporal, and multimedia data, as well as the fact that different types of data require different query capabilities [30]. Adding explicit support for a special kind of data is something that is beyond the absolute core of a standard, i.e., it can be seen as adding a cross-domain main model according to Figure 2 below the meta-model (which is part of the absolute core). As most IoT and DT standards considered in this paper are either rather new or still being defined, they are still focusing on the core of the standards and often did not yet add explicit support for different kinds of data. In this paper, we, therefore, consider only the support for geo-spatial, historical, and temporal data as these are the most important aspect that are already covered at least by some platforms.

Geo-spatial data is typically needed when there is any interaction with the physical world, e.g., a resource with a physical location, typically somewhere on earth but in some scenarios, this could also be somewhere in space. Historical and temporal data is quite similar, as both require data to be considered in the context of time. Historical data represent the history of a resource over time with a focus on the timely order of different states. Temporal data is considered simply data in the context of time. It can comprise data about the past but also the future of different resources at once, e.g., time-series data or results of prediction algorithms.

If standards support these different kinds of data is often determined by the paradigm followed. Probably most common is a device-centered paradigm where the virtual resource representing a physical device is considered a proxy reflecting the current state of the physical device as-is. Therefore, historical and temporal data is typically not available following this paradigm. Less common is an observation-based paradigm where observed changes in the environment are not only represented as changes in the state but also documented as an observation that was made by a sensor.

2.1.8. Resource Interlinking

As IoT and DT are highly dynamic environments of heterogeneous devices/resources, interlinking of resources essential to represent their (network) structure. The considered standards have different approaches and levels of support for resource interlinking which will be analyzed in detail for each standard in Section 3.

2.1.9. Semantic Annotation

In general, semantics is defined as 'the meaning or relationship of meanings of a sign or set of signs' [31]. In this context, signs are elements such as classes, properties, relations or instances of a knowledge base, expressed by their identifiers, e.g., http://qudt.org/vocab/unit/#DegreeCelsius. In technical systems, multiple levels of interoperability are to be considered [32]. Technical interoperability enables exchanging data on a physical level, syntactic interoperability that a system can parse the incoming data and semantic interoperability refers to 'the ability of computer systems to exchange data with unambiguous, shared meaning' [33]. As application models are not part of the standards, it cannot be expected that they are known to a consumer. Semantic annotation allows including references in the used (application) model(s), enabling consumers to discover the meaning of classes and properties on the fly.

2.2. Resource Discovery

For a consumer to know about existing resources, a discovery mechanism is needed. Discovering resources connected to a network is a common task, e.g., billions of resources are discovered via search engines every day.

Multiple well-established paradigms addressing resource discovery already exist, e.g., directories-style systems, search engines, crawlers, networking protocols (such as UPnP), etc. They can be divided into two groups; the ones defining how information about the existing resource is gathered and the ones defining how this gathered information could be accessed by users or other software systems. Information gathering can happen via e.g., (manual) registration of resources, networking protocols or crawlers. All of the considered standards default to manual or self-registration of a resource with some kind of resource directory. Some of them consider using special network protocols for resource discovery in the future but currently do not provide any solutions in this aspect. Therefore, we focus on how to access gathered information about existing resources, especially with the focus on supported network protocols and query capabilities.

2.2.1. Communication Protocols

Which communication protocols are supported for communication with the resource directory?

2.2.2. Query Capabilities

When trying to access a resource, a consumer typically does not want to access just any resource but a resource matching some defined criteria, e.g., measuring a certain property or offering a certain service. Although not always supported, it is a good idea for a resource directory to allow searching for resource matching certain criteria. In case searching is not supported, resource directories typically return the complete list of known resources, which then has to be filtered client-side. Therefore, our first classification criterion is whether queries generally are supported by the system.

In this paper, we focus only on the, from our perspective, most important aspects of query languages which are explained in the following. As query languages are often rather complex and differ in many small details, this set of classification criteria could be extended in the future. Because already many query languages exist, the most important criterion for us is which pre-existing query languages are supported. Typically, one pre-existing query language is chosen and used in a standard but some standards also support more than one or create a new one. Another criterion is if the query language(s) supports geo-spatial and temporal/historical queries. Although not directly affecting the query capabilities, we consider the possibility to add information about desired result formatting in a query a relevant criterion as it can massively improve usability and reduce network traffic when done right.

2.3. Resource Access

The actual goal in IoT and DT is accessing resources and communicating with them. Resource description and discovery are only intermediate steps required to achieve this goal. Of course, accessing and communicating with resources without underlying (IoT and DT) standards is also possible. However, this inevitably leads to writing lots of 'glue code', i.e., manually written code to communicate with different types of devices as they all use their own data formats and protocols. Especially in such a heterogeneous and dynamic environment like IoT where new devices are developed, connected to the internet and dynamically discovered by applications, writing clue code no longer is a viable solution.

2.3.1. Paradigm

We identified two different paradigms regarding resource access: API definition and API description. The most common paradigm to ensure unified resource access is to define some kind of resource API and require every resource to implement it. Although this is the easiest and most common approach, it does not take into account resources with existing APIs that cannot be updated to expose a new API. To properly address these issues and allow integration of legacy devices, the other paradigm is to not define a new resource API but rather provide a meta-model to describe existing APIs. On the downside, a lot of complexity is shifted to the client, i.e., implementing a client library becomes much harder as they now have to not only execute predefined commands via a single predefined protocol, but also understand the API description and potentially speak multiple protocols to communicate with a resource.

2.3.2. Communication Protocols

Resource access can happen via multiple different protocols, such as HTTP, CoAP, MQTT or OPC UA. Besides that, standards can be defined in a protocol-agnostic way and allow for mappings to other protocols to be added dynamically. This is especially important for standards using the API description paradigm.

3. Standards for Digital Twins and the Internet of Things

In this section, we introduce different standards related to TD and IoT and evaluate them according to the classification criteria defined in the previous section. The selection of standards was done based on discussions in different standardization bodies [34–36], recent scientific publications, as well as the authors' knowledge and experience in the areas of DT and IoT and their subjective relevance of the standards in practice. A previous work [5] focused on the extension of DTs with cognitive capabilities provides some additional insights to existing IoT and DT implementations.

3.1. Asset Administration Shell

The Asset Administration Shell (AAS) is a concept developed by the Plattform Industrie 4.0, a network of companies, associations, trade unions, science and politics in Germany [37]. It is part of the Reference Architectural Model Industry 4.0 (RAMI4.0) [38] and driven by the idea to manifest the concept of DT in factories and industrial production plants. The standard is divided into three parts. Part 1 [39] introduces the AAS meta-model and specifies how to serialize the content of an AAS, Part 2 defines how to interact with an AAS instance at runtime and Part 3 covers the infrastructure aspects and how to discover and interconnect multiple AASs [40]. This matches exactly our identified basic set of common functions: resource description, resource discovery, and resource access. At the time of writing, only Part 1 is officially released. Part 2 is available to the authors in a draft version.

An AAS is, just like a DT, a digital representation of a resource. Resources are referred to by the term 'asset', hence the name AAS. Figure 3 shows a class diagram of the AAS meta model. They main element is the AssetAdministrationShell, representing an Asset and composed of multiple Submodels. A Submodel is a container for SubmodelElements that can be properties, operations, collection, files, etc. The intention behind grouping sets of elements together into submodels and not adding those directly to an AAS is to allow standardization of submodels. By supporting multiple types of resource identifiers, IRI, IRDI and custom types, the standard provides much flexibility when referring to external resources e.g., for semantic annotations. The used type system is adapted from XSD complemented by the RDF data type langString. AAS descriptions can be serialized using XML, JSON, RDF, OPC UA data model and AutomationML. There is no explicit support for neither geo-spatial, temporal nor historical data although the standard is generic enough so this could be implemented on a higher level.

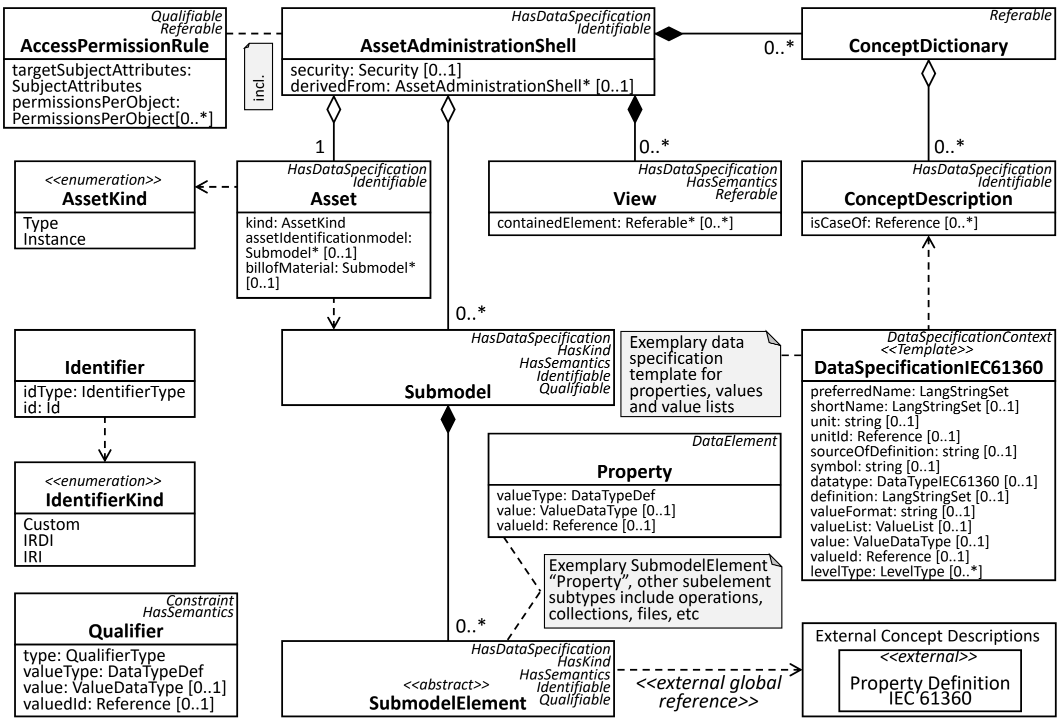


Figure 3. Class diagram depicting the metamodel of the AssetAdministrationShell concept developed by Plattform Industrie 4.0 (based on Figure 11 from [39]).

Resource discovery for this standard cannot be evaluated, as Part 3 is not yet released. However, Part 2 and [40] indicate that there will be an AAS registry to register AAS instances with some level of liability and a minimum set of metadata, e.g., endpoints. As an extension, AAS directories or even AAS catalogs may be built with further meta data elements that can be queried to discover resources.

For resource access, the standard will define a set of technology-specific APIs that each AAS instance has to implement. These APIs are derived from conceptual, technology-neutral interface specifications, which are currently being defined in Part 2. For the design of Industrie 4.0 systems and applications, they will be aggregated into services [41] by system and software architects, optionally extended with other, possibly proprietary interfaces. Therefore, the following information may be subject to changes in the future. At the core of the interfaces are basic CRUD (create, read, update, delete) operations for submodels and submodel elements. The interfaces will be defined for multiple communication protocols: HTTP(S), OPC UA, and probably MQTT.

3.2. Digital Twin Definition Language

The Digital Twin Definition Language (DTDL) [42] is developed by Microsoft and used in different products of their Azure services. Although not developed by an SDO, the DTDL is mentioned here because being already used in many commercial services offered by Microsoft like IoT Hub, IoT Central, and Azure Digital Twins. As the name suggests, the DTDL only covers the aspect of resource description and does not address resource discovery and resource access. There are two versions: Version 1 and Version 2. If not otherwise mentioned, we refer to Version 2 in this paper.

In the DTDL, resources are called interfaces and can contain a set of telemetry, properties, commands, relationship, and components as shown in Figure 4. For identification of resources and their elements DTDL uses a special form of URIs called Digital Twin Modeling Identifier (DTMI) of the form *<scheme>:<path>;<version>*. Telemetry describes data emitted by a resource in the form of a data stream. It corresponds to a combination of events and property subscriptions in our general terms. Commands correspond to functions that can be invoked with optional input and output parameters. Components are a similar concept as the submodels in the AAS providing a way to structure functionality in re-usable blocks. DTDL uses a custom type schema called Digital Twin Schema Definition Language, which is compatible with popular serialization formats, including JSON, Avro, and Protobuf. It supports complex interlinking of resources through explicit modeling of relationships including min/max cardinality. DTDL also supports semantic annotation for the type and unit of properties and telemetries. Unfortunately, only predefined semantic annotations are supported, i.e., extensibility is not given. JSON and RDF are supported in serialization formats. DTDL does neither support geo-spatial, temporal, nor historical data. Although in Version 1 a schema for geo-spatial data is defined (based on GeoJSON), it is not mentioned in Version 2 and therefore remains unclear if it still applies.

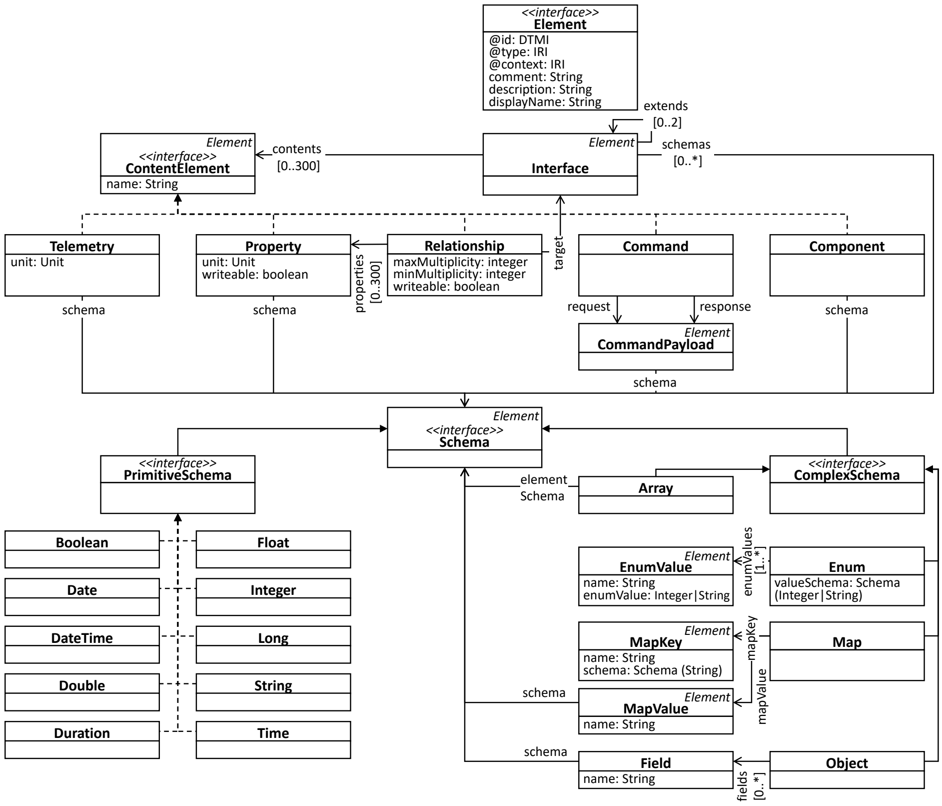


Figure 4. Class diagram depicting the metamodel of the digital twin definition language.

3.3. Next Generation Service Interfaces-Linked Data API

The Next Generation Service Interfaces-Linked Data (NGSI-LD) API [43] is an IoT standard defined by the ETSI Industry Specification Group (ISG) crosscutting Context Information Management (CIM). I is based on the Open Mobile Alliance (OMA) NGSI 9 and 10 interfaces [44] and FIWARE NGSIv2 [45]. In this paper, we refer to version 1.2.2 released 02/2020 [43]. Although never mentioning the term DT, the standard revolves around so-called 'Entities represent[ing] physical or conceptual objects existing in the real world' [46], which is more or less the definition of DT. In contrast to other standards, resources (entities) in NGSI-LD only comprise properties and relations to other resources. Services and events are not supported, although events can be partially realized through subscribing to property changes, which is supported.

As shown in Figure 5, NGSI-LD does not only provide a meta model defining general concepts such as Entity, Relationship and Property, but also a cross-domain ontology defining essential temporal and geo-spatial terms and concepts. Besides explicitly providing temporal and geo-spatial terms and concepts, the standard further supports historical data by allowing querying the state of an entity at any time in the past. Although, on a technical level, NGSI-LD could support resource identification via IRIs, they decided to allow only URIs (URI $\subset$ IRI) for the sake of being compliant to the Linked Data principles [47]. Because of the focus on Semantic Web technologies and Linked Data principles, NGSI-LD provides good support for semantic annotations. Serialization happens in JSON-LD, which since version 1.1 is compatible to basic JSON as well as RDF.

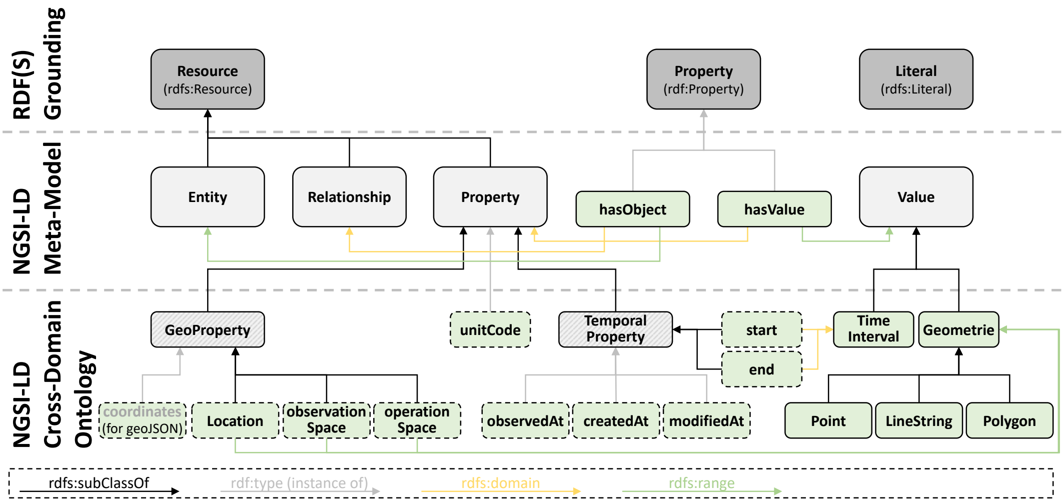


Figure 5. Class diagram depicting the metamodel hierarchy of the Next Generation Service Interfaces-Linked Data (NGSI-LD) API as defined by European Telecommunications Standards Institute (ETSI) (based on Figure 4.2.3-1 from [43]).

Resource discovery is realized by a component named 'Context Registry'. NGSI-LD supports different types of architectures: centralized, distributed, and federated. In the centralized architecture, only one context registry exists whereas in the distributed and federated there can be multiple (hierarchically organized) registries. Any registry supports resource discovery via HTTP and custom query languages enabling geo-spatial, temporal and historical queries. Although there are already well-established query languages for RDF such as SPARQL or GeoSPARQL, NGSI-LD decided to define custom query languages from scratch. The reason for this is that NGSI-LD does not use the RDF data format where data is represented as triples in the form of <subject, predicate, object> but rather uses the concept of property graphs [48] where predicates of such triples can contain further key-value data.

Resource access in NGSI-LD is fundamentally different from the other standards as there is no direct communication between producer and consumer. Instead, the producer provides their information to a broker and consumers can then access these data or subscribe to be notified upon changes. These interactions are defined in a communication protocol-agnostic manner but currently only a binding to HTTP is provided.

3.4. Open Data Protocol

The Open Data Protocol (OData) [49] allows the definition of standardized but data model-agnostic RESTful APIs. It was initially developed by Microsoft in 2007 and Version 1.0, 2.0, and 3.0 were released under the Microsoft Open Specification Promise. In 2014, Version 4.0 was standardized by the Organization for the Advancement of Structured Information Standards (OASIS). This is the version we are referring to in this paper. Although not developed with IoT or DT explicitly in mind, we consider OData relevant for them as it is all about describing and accessing resources (even though originally resources were not considered physical).

The primary concept of OData is called 'service'. Amongst others, a service comprises entity and entity set definitions as well as functions (must not have side effects, must have a result) and actions (may have side effects, may have a result). OData services are described by an Entity Data Model (EDM) expressed via the Common Schema Definition Language (CSDL), which in turn can be serialized via XML and JSON. The CSDL, therefore, corresponds to what we refer to as resource description in this article. The identification of resources happens primarily via URLs. Additionally, resources can also be identified via their entity type and ID, which, together with the URL conventions, implicitly defines the URL [50]. Resources can easily be interlinked and contain annotations, which can be used to add (external) semantic. OData do not support defining events on resources, nor subscribing to property changes.

Similar to NGSI-LD, there is no direct communication between producer and consumer in OData and all communication happens via a central server running the service. The standard defines a HTTP-based API for communication with the service that includes a powerful custom URL-based query language supporting geo-spatial filtering.

As all communication, resource access in OData happens via HTTP(S). URLs in OData are defined by the URL conventions in combination with the data model defined by the EDM. An example URL is shown in Figure 6. The service root part is given by the OData service and can be followed by any entity set name in the resource path. After that, any path defined by the EDM is allowed, e.g., accessing a single entity in an entity set by appending its ID in parenthesis as well as following relations or addressing properties by adding a followed by the name. The query options part of a URL can be used to specify the desired result more fine-grained.

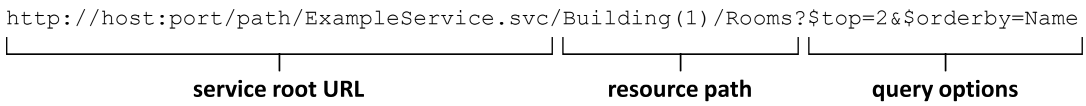


Figure 6. Components of URLs according to the Open Data Protocol specification.

3.5. SensorThings API

The SensorThings API (STA) standard [51] developed by the OGC SensorThings Standards Working Group (SWG) [52] provides a framework for interconnecting IoT devices, data, and applications over the Web. It is divided into multiple parts. Part 1 [53] was published in 2016 and covers the domain of sensing. Part 2 [54] addresses tasking, i.e., sending commands to devices, and was published in 2019. An updated version (v1.1) of Part 1 is finished and formal publication is underway.

Since 2001, the OGC is working on a suite of standards for managing sensors and actuators called OGC Sensor Web Enablement (SWE). It is designed to enable the connecting sensor and actuator to the Web and make them discoverable and query-able. As sensors and actuators are usually physical devices, geo-location is a key element of OGS standards. The older SWE standards, e.g., Sensor Observation Service (SOS) and Sensor Planning Service (SPS), use rather old technologies like XML and SOAP but more importantly, they are also missing some basic features like publish/subscribe (pub/sub) support or capabilities to update/delete data. STA addresses those issues and at the same time redesigns the APIs to fit modern web-based principles like REST-based services and JSON as payload format.

STA is strongly inspired by OData and re-uses some essential aspects, e.g., most parts of the URL conventions and query language, but is not fully compliant to OData. Besides leaving out some complexity of OData, STA also adds functionality where needed, e.g., pub/sub via MQTT and additional geo-spatial query capabilities. As OData itself is model-agnostic, STA provides a cross-domain model that can be further refined for any use case by providing an application model according to the metamodel hierarchy shown in Figure 2. The STA data model is shown in Figure 7. In STA, resources are called Things and are identified using URLs. The type system used is custom-built but re-using multiple existing type definitions, e.g., from previous SWE standards. Services are realized by modeling (virtual or physical) actuators with their capabilities and linking them to a thing. Events are only supported in the form of a subscription of changes of properties via MQTT. Resource interlinking and semantic annotation is hardy explicitly supported (only the class ObservedProperty has a property definition for semantic annotation) but can potentially be added by custom extension. Historical data is also only partially supported, as observations are stored over time, i.e., discovering past observations are possible but reading the value of properties at any time in the past is not possible.

Resource discovery works almost identically as in OData although STA provides enhanced geo-spatial query capabilities.

Resource access works also very similar to OData as both do not support direct communication between resource and consumer. In addition to OData, STA allows subscribing to the creation and update of resource collections, resources and properties. Having no direct communication between a consumer and a resource drastically improves usability in real-world applications, e.g., an actuator can subscribe to newly created tasks and immediately start processing them.

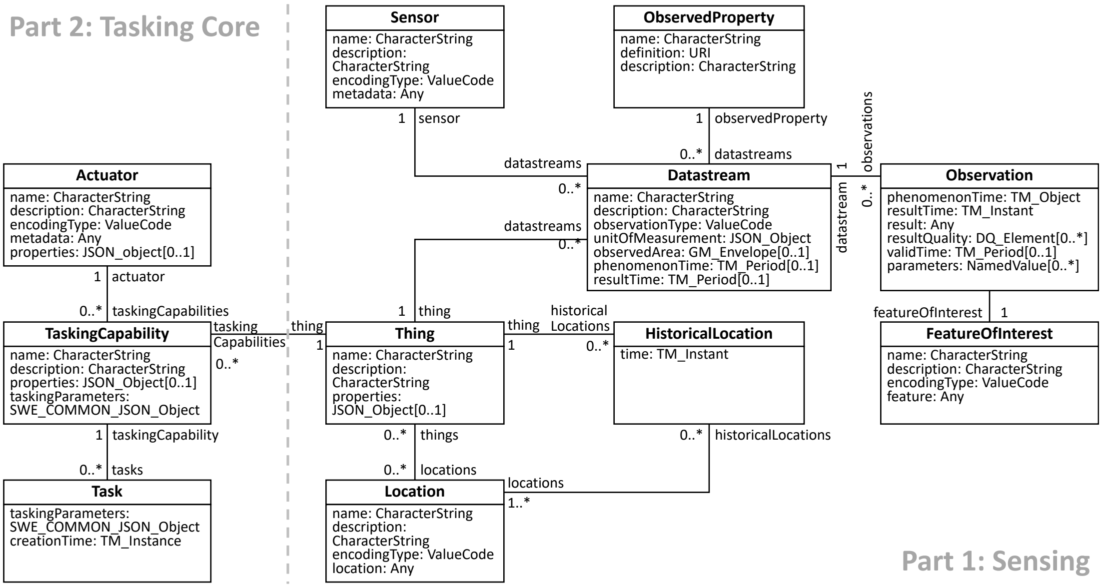


Figure 7. Class diagram depicting the data model of the SensorThings API Part 1 and 2 in version 1.0 (based on Figure 2 from [53] and Figure 1 from [54]).

3.6. Web of Things

The goal of the W3C Web of Things (WoT) Working Group is to counter the fragmentation of the IoT by providing so-called building blocks that should complete and enhance already existing standards. The most important of those building blocks is the WoT ThingDescription (WoT TD or only TD) [55], which became an official W3C Recommendation in April 2020. The TD provides a meta-model to describe existing resources in the IoT, e.g., a DT or a part of a DT, called Thing. As Figure 8 shows, a Thing comprises a set of properties (which can be read-only or read/write), actions, i.e., methods that can be invoked on the Thing, and events. The type system is based on JSON with additional support for adding constraints, e.g., regular expression for strings, based on JSON Schema. TDs use URI-based identifiers and can be serialized as either JSON-LD, JSON, or RDF. The standard does not contain any kind of (cross-)domain model and therefore neither supports geo-spatial, temporal nor historical data. One highlight of the TD is that by being grounded in Semantic Web technologies, is it easily possible to add semantic information to every aspect of a TD.

A serialized TD can be published and a consumer will be able to access the described resource. Resource discovery, i.e., how these TDs can be found, is still work in progress but is expected to be published in Q1 2021 [56]. Not yet standardized, there are also at least two existing implementations (https://github.com/thingweb/thingweb-directory) (https://github.com/linksmart/thing-directory) providing sophisticated search capabilities, e.g., via DNS-Based Service Discovery [57], CoRE Resource Directory [58], SPARQL queries or JSON-LD framing [59].

Unlike other standards, the TD does not define any API that has to be implemented by any resource. Instead, it defines a meta-model to describe existing APIs including e.g., protocol, payload format, security, in a machine-readable way. This is a unique and promising approach at closing the gap between existing (and future) standards and enabling interoperability.

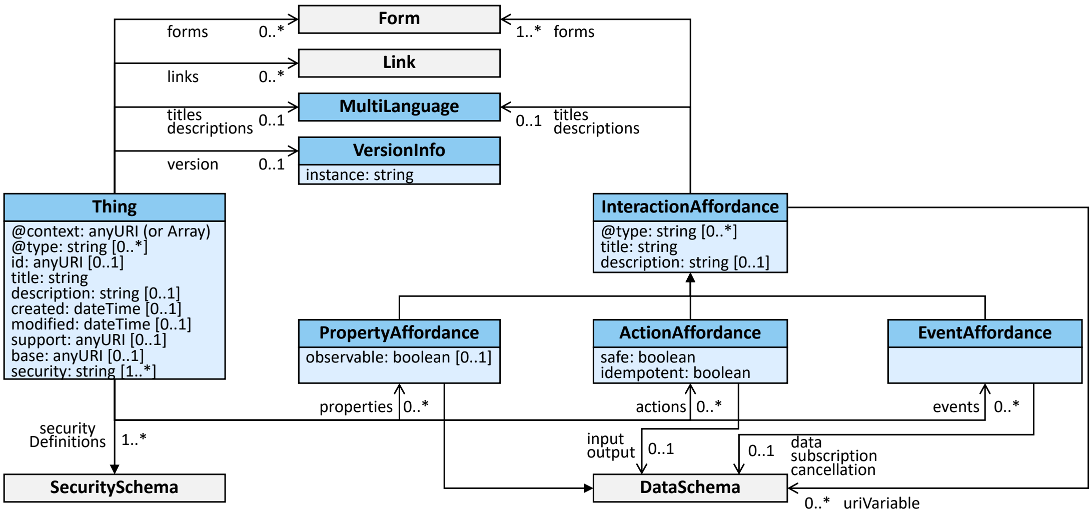


Figure 8. Class diagram depicting the metamodel of the Web of Things Thing Description (based on Figure 1 from [55]).

4. Summary

Table 1 provides an overview on the classification of the different IoT and DT standards analyzed. Although the standards were developed independently by different SDOs, they ended up with quite similar approaches/solutions to some aspects.

To enable resource description, providing a meta-model is the way to go. NGSI-LD and STA also provide some cross-domain model, primarily defining terms needed for support of geo-spatial and temporal data. At first glance, the standards seem to use quite different types of resource identifiers. Considering that IRI $\supset$ URI $\supset$ URL, URI $\supset$ DTMI, and URL $\cap$ DTMI = $\varnothing$ it seems that URI might be a good compromise for a unified identifier type in the future. An even more flexible, but also more complex, approach is the one used in the AAS by defining meta-level constructs to describe any kind of identifier. All standards supporting resource interlinking (or at least having it on the agenda) reflects the fact that relations between resources are essential for IoT and DT. Besides, using different terms to describe the kind of elements that make up a resource, there seems to be more-or-less general consent that a resource should at least contain properties and services. Only half of the standards support resources defining custom events, but almost all support at least notification upon property changes. For serialization of resource descriptions, JSON is the common denominator between all standards. Since version 1.1, JSON-LD [60] supports interpreting basic JSON as JSON-LD when using HTTP and setting the header appropriately. Additionally, JSON-LD is compatible with RDF, the cornerstone of the Semantic Web. For this reason, almost all standards make use of JSON-LD to provide, on the one hand, easy-to-use JSON-based access and, on the other side, fine-grained access with full Semantic Web support. Again, the AAS provides a more generic approach supporting different kinds of serialization formats. This is partially because it also supports different protocols for resource access and some protocols require a certain serialization format, e.g., OPC UA.

Resource discovery is something on the agenda of most of the SDOs and will be properly addressed in the future. No standards currently supporting resource discovery allow direct communication between the consumer and the resource, i.e., they already have a central repository-like component by design. Furthermore, there is no consensus about which query language to use. Query languages are a complex subject itself. In fact, AAS and WoT are currently still evaluating different query languages that might be suited for them. Probably the main issue is the trade-off between the expressivity of a language and the complexity to implement it. Additionally, query languages are also often tied to specific data formats and/or protocols. SPARQL is a rather complex but powerful query language for RDF with SQL-like syntax. As most standards already support RDF data, SPARQL is one of the most promising candidates for a unified query language. Furthermore, there are already existing extensions for geo-spatial queries in SPARQL, e.g., GeoSPARQL. Another promising approach are URL-based query languages like OData/STA and JSONPath. The query language of OData/STA has proven quite powerful in real-world applications but is much simpler to implement than SPARQL. JSONPath is adapting the concept of XPath from XML to JSON. It has been around for over ten years but has never been properly standardized by an SDO. Recently, efforts have been made initiated by the IRTF Thing-to-Thing Research Group (T2TRG) to pursue an official standardization and to push JSONPath as a query language in the IoT context [34,61].

Table 1. Comparison of the following internet of things and digital twin standards: AssetAdministrationShell (AAS), Digital Twin Definition Language (DTDL) v2.0, Next Generation Service Interfaces—Linked Data API (NGSI-LD API), Open Data Protocol (OData) v4.0, SensorThings API (STA) v1.0, and Web of Things (WoT).

	AAS	**DTDL**	**NGSI-LD**	**OData**	**STA**	**WoT**
Resource Description						
Resource Term	Asset	Interface	Entity	Entity	Thing	Thing
Model Type(s)	Meta	Meta	Meta Cross-Domain	Meta	Cross-Domain	Meta
Resource Identification	IRI IRDI custom	DTMI	URI	URL custom	URL custom	URI
Type System (based on)	XSD	custom	JSON GeoJSON JSON-LD	custom	JSON SWE-standards	JSON JSON Schema
Resource Interlinking	X	X	X	X	- [a]	X
Semantic Annotation	X	O [b]	X	-	O [c]	X
Resource Elements						
Properties	X	X	X	X	X	X
Services	X	X	-	O [d]	O	X
Events	X	X	O [e]	-	O [e]	X
Serialization Format	JSON RDF XML OPC UA AutomationML	JSON RDF Avro Protobuf	JSON RDF	JSON XML	JSON	JSON RDF
Supported Kind of Data						
geo-spatial	-	-	X	X	X	-
temporal	-	-	X	X	X	-
historical	-	-	X	-	O [f]	-
Resource Discovery						
Protocols	- [a]	-	HTTP	HTTP	HTTP	HTTP [g] CoAP [g] DNS-SD [g]
Querying supported?	- [a]	-	X	X	X	O [g]
Query Language						
Query Language based on	- [a]	-	custom	custom	OData	SPARQL [a,g]
geo-spatial queries	-	-	X	X	X	-
historical queries	-	-	X	-	O [f]	-
Resource Access						
API: Define vs. Describe	define	-	define	define	define	describe
Protocols	HTTP MQTT OPC UA	-	HTTP	HTTP	HTTP MQTT	HTTP MQTT CoAP
Protocols extendible?	-	-	X	-	-	X

[a] extension under discussion; [b] only predefined definitions and only for telemetries, properties, and units; [c] only explicitly for observed properties and units, possible for everything else via custom properties; [d] only on service-level; [e] only property changes; [f] only for observations; [g] not part of standard, only in implementation(s); [x] y; Abbreviations: CoAP: Constrained Application Protocol; DNS-SD: Domain Name System - Service Discovery; HTTP: Hypertext Transfer Protocol; IRDI: International Registration Data Identifier; IRI: Internationalized Resource Identifier; JSON: JavaScript Object Notation; JSON-LD: JavaScript Object Notation - Linked Data; MQTT: Message Queuing Telemetry Transport; OPC UA: Open Platform Communications Unified Architecture; RDF: Resource Description Format; SPARQL: SPARQL and RDF Query Language; SWE: Sensor Web Enablement; URI: Uniform Resource Identifier; URL: Uniform Resource Locator; XML: Extensible Markup Language; XSD: XML Schema Definition.

Regarding resource access, the most prominent paradigm is to define a new API that each resource has to implement. The only standard adapting to another paradigm is WoT. Their idea is that there are already plenty of IoT devices and systems deployed and changing/updating their interfaces is not an option. Therefore, the WoT approach is to provide a way to describe existing resources and their APIs. Although WoT is the only standard following this approach, it seems like a valid claim considering there are already trillions of IoT devices deployed, often without the possibility to update them over-the-air. For resource access protocols, HTTP seems to be general consent. Typically, it is accompanied by MQTT adding pub/sub functionality that HTTP is missing. Not limiting itself to a specific set of communication protocols seems like a good idea, as APIs should, in general, be protocol agnostic and also because the past shows that almost every communication protocol evolves over time and will eventually be replaced.

In general, it is to notice, that besides IoT and DT being broad domains with heterogeneous applications and partially different requirements, there exists some shared common core functionality of resource description, discovery and access. Different standards, developed by different SDOs, having different backgrounds like, e.g., environmental measurements, industrial production, appliances, or consumer electronics, came up with partially very similar standards, showing that there might be a chance for even better convergence. Furthermore, there is a need to abstract from the underlying technologies when mapping requirements on resource management and their DT counterparts systematically to capabilities of underlying IoT infrastructures [62]. DT Engineering will become a crucial methodology in future IoT environments.

5. Conclusions

The paper highlighted the relevance of IoT and DT and the importance of interoperability and standardization for industry 4.0. The description and managing of resources were identified as overlapping functionality between DT and IoT that is addressed by standards from both domains that need consolidation. This paper introduced a classification scheme based on the categories resource description, discovery, and access and presented relevant current standards from both domains. The standards were evaluated using the classification scheme and the results were compared. This analysis revealed commonalities and differences between the considered standards that can be used to stimulate the consolidation of standards in the future. Commonalities can essentially be found for basic aspects, e.g., that resources should be described by a set of properties, services, and events and that JSON-LD should be used as a serialization format for resource descriptions. There further seems to be an agreement that resource access through HTTP should be the default while additional protocols can be supported. Differences are typically related to more complex aspects, such as if geo-spatial, temporal, and historical data should be explicitly supported or which query language to use for resource discovery.

Some standards, especially the ones focused on DTs, are divided into multiple parts, and can be considered work in progress, as not all of those parts are published yet. The problem of competing and overlapping standards in these domains and the need for consolidation has already been acknowledged [35,63]. Cooperation between the involved SDOs is already happening but primarily on a working group level in a rather informal manner, e.g., by inviting representatives of other SDOs to present their work or persons that are active in multiple relevant working groups. This paper may be used as a starting point to steer this cooperation to a more formalized and technical level and thereby stimulate and advance the consolidation of standards in these domains.

In future work, the presented classification scheme can be extended and elaborated, e.g., with focus on semantic interoperability. Furthermore, additional standards may be added or existing ones updated as they evolve over time.

Author Contributions: M.J.: conceptualization, methodology, investigation, writing—original draft preparation, and visualization; T.U.: writing—review and supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partly funded by the H2020 COGNITWIN project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 870130.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AAS	Asset Administration Shell
AI	Artificial Intelligence
API	Application Programming Interface
BDVA	Big Data Value Association
CIM	Context Information Management
CoAP	Constrained Application Protocol
CoRE	Constrained RESTful Environments
CRUD	Create, Read, Update, Delete
CSDL	Common Schema Definition Language
CT	Cognitive Twin
DNS	Domain Name System
DT	Digital Twin
DTDL	Digital Twin Definition Language
DTMI	Digital Twin Model Identifier
EDM	Entity Data Model
ETSI	European Telecommunications Standards Institute
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure
I4.0	Industry 4.0
IEC	International Electrotechnical Commission
IETF	Internet Engineering Task Force
IIC	Industrial Internet Consortium
ISG	Industry Specification Group
IoT	Internet of Things
IoT-EPI	IoT-European Platforms Initiative
IRDI	International Registration Data Identifier
IRI	Internationalized Resource Identifier
IRTF	Internet Research Task Force
ISO	International Organization fpr Standardization
ITU-T	International Telecommunication Union Telecommunication Standardization Sector
JSON	JavaScript Object Notation
LD	Linked Data
M2M	machine-to-machine
ML	Machine Learning
MQTT	Message Queuing Telemetry Transport
NASA	National Aeronautics and Space Administration
NGSI	Next Generation Service Interfaces
OASIS	Organization for the Advancement of Structured Information Standards
OData	Open Data Protocol
OGC	Open Geospatial Consortium
OMA	Open Mobile Alliance
OMG	Object Management Group
OPC UA	Open Platform Communications Unified Architecture
OWL	Web Ontology Language
pub/sub	publish/subscribe
RAMI4.0	Reference Architectural Model Industry 4.0
RDF	Resource Description Format
RDFS	RDF Schema
REST	Representational State Transfer
SDO	Standards Developing Organization
SOAP	Simple Object Access protocol
SOS	Sensor Observation Service
SPARQL	SPARQL and RDF Query Language
SPS	Sensor Planning Service
SSN	Semantic Sensor Network (Ontology)

STA	SensorThings API
SWE	Sensor Web Enablement
SWG	Standards Working Group
T2TRG	Thing-to-Thing Research Group
TD	ThingDescription
UPnP	Universal Plug and Play
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
XML	Extensible Markup Language
XSD	XML Schema Definition
W3C	World Wide Web Consortium
WoT	Web of Things

References

1. Forbes. What Is Industry 4.0? Here's a Super Easy Explanation for Anyone. 2018. Available online: https://www.forbes.com/sites/bernardmarr/2018/09/02/what-is-industry-4-0-heres-a-super-easy-explanation-for-anyone/ (accessed on 9 August 2020).
2. Plattform Industrie 4.0. What Is Industrie 4.0? Available online: https://www.plattform-i40.de/PI40/Navigation/EN/Industrie40/WhatIsIndustrie40/what-is-industrie40.html (accessed on 9 August 2020).
3. Hermann, M.; Pentek, T.; Otto, B. Design principles for industrie 4.0 scenarios. In Proceedings of the 49th Hawaii international conference on system sciences (HICSS), Koloa, HI, USA, 5–8 January 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 3928–3937.
4. Malakuti, S.; van Schalkwyk, P.; Boss, B.; Sastry, C.R.; Runkana, V.; Lin, S.W.; Rix, S.; Green, G.; Baechle, K.; Nath, S.V. Digital Twins for Industrial Applications: Definition, Business Values, Design Aspects, Standards and Use Cases. In *Industrial Internet Consortium White Paper*; Industrial Internet Consortium: Milford, MA, USA, 2020.
5. Abburu, S.; Roman, D.; Berre, A.; Stojanovic, L.; Jacoby, M.; Stojanovic, N. COGNITWIN–Hybrid and Cognitive Digital Twins for the Process Industry. In Proceedings of the IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), 2020, Forthcoming.
6. Digital Twin Market Worth $48.2 Billion by 2026. Available online: https://www.marketsandmarkets.com/PressReleases/digital-twin.asp (accessed on 9 August 2020).
7. Deloitte. Industry 4.0: Challenges and Solutions for the Digital Transformation and Use of Exponential Technologies. Available online: https://www2.deloitte.com/content/dam/Deloitte/ch/Documents/manufacturing/ch-en-manufacturing-industry-4-0-24102014.pdf (accessed on 9 August 2020).
8. Schleich, B.; Dittrich, M.A.; Clausmeyer, T.; Damgrave, R.; Erkoyuncu, J.A.; Haefner, B.; de Lange, J.; Plakhotnik, D.; Scheidel, W.; Wuest, T. Shifting value stream patterns along the product lifecycle with digital twins. *Procedia CIRP* **2019**, *86*, 3–11.
9. Tao, F.; Zhang, M.; Nee, A.Y.C. *Digital Twin Driven Smart Manufacturing*; Academic Press: Cambridge, MA, USA, 2019.
10. Gartner, Inc. The Top 10 Technology Trends for 2012. Available online: https://www.gartner.com/en/documents/1926316/the-top-10-technology-trends-for-2012 (accessed on 9 August 2020).
11. Gartner, Inc. The Top 10 Strategic Technology Trends for 2013. Available online: https://www.gartner.com/en/documents/2335015/the-top-10-strategic-technology-trends-for-2013 (accessed on 9 August 2020).
12. Gartner, Inc. The Top 10 Strategic Technology Trends for 2014. Available online: https://www.gartner.com/en/documents/2667526/the-top-10-strategic-technology-trends-for-2014 (accessed on 9 August 2020).
13. Gartner, Inc. The Top 10 Strategic Technology Trends for 2015. Available online: https://www.gartner.com/en/documents/2964518/the-top-10-strategic-technology-trends-for-2015 (accessed on 9 August 2020).
14. Rivera, J.; van der Meulen, R. Forecast Alert: Internet of Things—Endpoints and Associated Services. 2017. Available online: https://www.gartner.com/en/documents/3559634/forecast-alert-internet-of-things-endpoints-and-associat (accessed on 9 August 2020).
15. Gartner, Inc. Top 10 Strategic Technology Trends for 2017. Available online: https://www.gartner.com/en/documents/3471559/top-10-strategic-technology-trends-for-2017 (accessed on 9 August 2020).

16. Gartner, Inc. Gartner Top 10 Strategic Technology Trends for 2018. Available online: https://www.gartner.com/en/documents/3867164/top-10-strategic-technology-trends-for-2018-digital-twin (accessed on 9 August 2020).
17. Gartner, Inc. Gartner Top 10 Strategic Technology Trends for 2019. Available online: https://www.gartner.com/en/documents/3904569/top-10-strategic-technology-trends-for-2019-digital-twin (accessed on 9 August 2020).
18. Gartner, Inc. Gartner Top 10 Strategic Technology Trends for 2020. Available online: https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2020 (accessed on 9 August 2020).
19. Gartner, Inc.. How Digital Twins Simplify the IoT. Available online: https://www.gartner.com/smarterwithgartner/how-digital-twins-simplify-the-iot/ (accessed on 9 August 2020).
20. Xu, L.D. The contribution of systems science to Industry 4.0. *Syst. Res. Behav. Sci.* **2020**, *37*, 618–631.
21. Burns, T.; Cosgrove, J.; Doyle, F. A Review of Interoperability Standards for Industry 4.0. *Procedia Manuf.* **2019**, *38*, 646–653.
22. Lu, Y. Industry 4.0: A survey on technologies, applications and open research issues. *J. Ind. Inf. Integr.* **2017**, *6*, 1–10.
23. Minerva, R.; Lee, G.M.; Crespi, N. Digital Twin in the IoT Context: A Survey on Technical Features, Scenarios, and Architectural Models. *Proc. IEEE* **2020**.
24. Global Digital Twins Market Insights 2020–2025 by Technology, Solution, Application and Industry Vertical. Available online: https://www.globenewswire.com/news-release/2020/03/20/2003898/0/en/Global-Digital-Twins-Market-Insights-2020-2025-by-Technology-Solution-Application-and-Industry-Vertical.html (accessed on 9 August 2020).
25. Transforma Insights. Global IoT Market Will Grow to 24.1 Billion Devices in 2030, Generating $1.5 Trillion Annual Revenue. Available online: https://transformainsights.com/news/iot-market-24-billion-usd15-trillion-revenue-2030 (accessed on 9 August 2020).
26. McKinsey Global Institute *The Internet of Things: Mapping the Value Beyond the Hype*. McKinsey&Company: New York, NY, USA, 2015.
27. Digital Twin in the Industry 4.0: Interview with A Pioneer. 2018. Available online: https://www.t-systems.com/de/en/about-t-systems/news/best-practice/03-2018-digital-twin/digital-twin-in-the-industry-4-0-interview-with-a-pioneer (accessed on 9 August 2020).
28. *Meta Object Facility (MOF) Specification, Version 1.4*; Object Management Group, Ed.; Object Management Group: Milford, MA, USA, 2002.
29. eCl@ss: Classification and Product Description. Available online: https://www.eclass.eu/ (accessed on 9 August 2020).
30. BDVA Strategic Research and Innovation Agenda (SRIA). 2017. Available online: https://bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf (accessed on 9 August 2020).
31. Merriam-Webster Dictionary: Semantics. Available online: https://www.merriam-webster.com/dictionary/semantics (accessed on 9 August 2020).
32. Wang, W.; Tolk, A.; Wang, W. The Levels of Conceptual Interoperability Model: Applying Systems Engineering Principles to M&S. In Proceedings of the Spring Simulation Multiconference, San Diego, CA, USA, 22–27 March 2009; p. 168.
33. Systems, Capabilities, Operations, Programs, and Enterprises (SCOPE) Model for Interoperability Assessment. Network-Centric Operations Industry Consortium, 2008. Available online: http://www.jfsowa.com/ikl/scope08.pdf (accessed on 9 August 2020).
34. IRTF Thing-to-Thing Research Group (T2TRG). Available online: https://datatracker.ietf.org/rg/t2trg/charter/ (accessed on 9 August 2020).
35. Reach Out to Other Open Standards. Available online: https://github.com/Azure/opendigitaltwins-dtdl/issues/10 (accessed on 9 August 2020).
36. Web of Things Working Group. Available online: https://www.w3.org/WoT/WG/ (accessed on 9 August 2020).
37. Details of the Asset Administration Shell: From Idea to Implementation. Plattform Industrie 4.0, Ed. Available online: https://www.plattform-i40.de/PI40/Redaktion/EN/Downloads/Publikation/vws-in-detail-presentation.pdf (accessed on 9 August 2020).

38. RAMI 4.0 - A Reference Framework for Digitalisation. Plattform Industrie 4.0, Ed. Available online: https://www.plattform-i40.de/PI40/Redaktion/EN/Downloads/Publikation/rami40-an-introduction.pdf (accessed on 9 August 2020).
39. Details of the Asset Administration Shell: Part 1-The Exchanges of Information between Partners in the Value Chain of Industrie 4.0 (Version 2.0.1). Plattform Industrie 4.0, Ed. 2020. Available online: https://www.plattform-i40.de/PI40/Redaktion/EN/Downloads/Publikation/Details-of-the-Asset-Administration-Shell-Part1.pdf (accessed on 9 August 2020).
40. Hoffmeister, M. Properties and the Asset Administration Shell as Information Backbone-How Can True Interoperability be Achieved along the Entire Industrie 4.0 Value Chain. Available online: https://www.eclass.eu/fileadmin/pdfs/1_16-55_Hoffmeister.pdf (accessed on 9 August 2020).
41. *DIN SPEC 16593-1:2018-04 RM-SA-Reference Model for Industrie 4.0 Service Architectures-Part 1: Basic Concepts of an Interaction-Based Architecture*; Usländer, T., Ed.; Beuth: Berlin, Germany, 2018. Available online: https://www.beuth.de/en/technical-rule/din-spec-16593-1/287632675 (accessed on 9 August 2020).
42. Digital Twins Definition Language. Available online: https://github.com/Azure/opendigitaltwins-dtdl/blob/master/DTDL/v2/dtdlv2.md (accessed on 9 August 2020).
43. ETSI GS CIM 009: NGSI-LD API. 2020. Available online: https://www.etsi.org/deliver/etsi_gs/CIM/001_099/009/01.02.02_60/gs_CIM009v010202p.pdf (accessed on 9 August 2020).
44. Next Generation Service Interfaces Architecture. Open Mobile Alliance, Ed. 2012. Available online: http://www.openmobilealliance.org/release/NGSI/V1_0-20120529-A/OMA-AD-NGSI-V1_0-20120529-A.pdf (accessed on 9 August 2020).
45. FIWARE-NGSI v2 Specification. Available online: https://www2.deloitte.com/content/dam/insights/us/articles/tech-trends-2020/DI_TechTrends2020.pdf (accessed on 9 August 2020).
46. ETSI GR CIM 008: Context Information Management (CIM); NGSI-LD Primer. 2020. Available online: https://www.etsi.org/deliver/etsi_gr/CIM/001_099/008/01.01.01_60/gr_CIM008v010101p.pdf (accessed on 9 August 2020).
47. Berners-Lee, T. Linked Data-Design Issues. 2006. Available online: https://www.w3.org/DesignIssues/LinkedData.html (accessed on 9 August 2020).
48. Defining a Property Graph. 2012. Available online: https://github.com/tinkerpop/gremlin/wiki/Defining-a-Property-Graph (accessed on 9 August 2020).
49. OData. Available online: https://www.odata.org/ (accessed on 9 August 2020).
50. OData Version 4.01. Part 2: URL Conventions. 2016. Available online: http://docs.oasis-open.org/odata/odata/v4.0/odata-v4.0-part2-url-conventions.pdf (accessed on 9 August 2020).
51. OGC SensorThings API. Available online: https://www.ogc.org/standards/sensorthings (accessed on 9 August 2020).
52. SensorThings SWG. Available online: https://www.ogc.org/projects/groups/sweiotswg (accessed on 9 August 2020).
53. OGC SensorThings API Part 1: Sensing, Version 1.0. Liang, S., Huang, C.-Y., Khalafbeigi, T., Eds.; Open Geospatial Consortium: Wayland, MA, USA, 2016. Available online: http://www.opengis.net/doc/is/sensorthings/1.0 (accessed on 9 August 2020).
54. OGC SensorThings API Part 2—Tasking Core, Verson 1.0. Liang, S., Khalafbeigi, T., Eds.; Open Geospatial Consortium: Wayland, MA, USA, 2019. Available online: http://www.opengis.net/doc/IS/sensorthings-part2-TaskingCore/1.0 (accessed on 9 August 2020).
55. Web of Things Thing Description. Available online: https://www.w3.org/TR/wot-thing-description/ (accessed on 9 August 2020).
56. Web of Things Working Group Charter. Available online: https://www.w3.org/2020/01/wot-wg-charter.html (accessed on 9 August 2020).
57. Cheshire, S.; Krochmal, M. RFC 6763: DNS-based Service Discovery. 2013. Available online: https://tools.ietf.org/html/rfc6763 (accessed on 9 August 2020).
58. Shelby, Z.; Koster, M.; Bormann, C.; van der Stok, P.; Amsuess, C. CoRE Resource Directory. 2017. Available online: https://tools.ietf.org/html/draft-ietf-core-resource-directory-12 (accessed on 9 August 2020).
59. Longley, D.; Sporny, M.; Kellogg, G.; Lanthaler, M.; Lindström, N. W3C Recommendation: JSON-LD 1.1 Framing. 2020. Available online: https://www.w3.org/TR/json-ld11-framing/ (accessed on 9 August 2020).

60. JSON-LD 1.1-A JSON-based Serialization for Linked Data. Available online: https://www.w3.org/TR/json-ld11/ (accessed on 9 August 2020).
61. Gössner, S.; Normann, C. JSONPath–XPath for JSON. 2020. Available online: https://tools.ietf.org/html/draft-goessner-dispatch-jsonpath-00 (accessed on 9 August 2020).
62. Usländer, T.; Batz, T. Agile service engineering in the industrial Internet of Things. *Future Internet* **2018**, *10*, 100. doi:10.3390/fi10100100.
63. W3C and IIC Sign Liaison Agreement for the Industrial Internet of Things. Available online: https://www.w3.org/blog/wotig/2015/08/11/w3c-and-iic-sign-liaison-agreement-for-the-industrial-internet-of-things/ (accessed on 9 August 2020).

applied sciences

Article

Implementation of Digital Twin for Engine Block Manufacturing Processes

Roman Bambura [1,*], Marek Šolc [2], Miroslav Dado [1] and Luboš Kotek [3]

1 Department of Manufacturing Technology and Quality Management, Faculty of Technology, Technical University in Zvolen, Študentská 26, 960 01 Zvolen, Slovakia; dado@tuzvo.sk
2 Institute of Materials and Quality Engineering, Faculty of Materials, Metallurgy and Recycling, Technical University of Košice, Park Komenského 3, 040 01 Košice, Slovakia; marek.solc@tuke.sk
3 Department of Production Systems and Virtual Reality, Faculty of Mechanical Engineering, Brno University of Technology, Technická 2896/2, 616 69 Brno, Czech Republic; kotek.l@fme.vutbr.cz
* Correspondence: xbambura@is.tuzvo.sk; Tel.: +421-45-520-6026

Received: 18 August 2020; Accepted: 17 September 2020; Published: 21 September 2020

Abstract: The digital twin (DT) is undergoing an increase in interest from both an academic and industrial perspective. Although many authors proposed and described various frameworks for DT implementation in the manufacturing industry context, there is an absence of real-life implementation studies reported in the available literature. The main aim of this paper is to demonstrate feasibility of the DT implementation under real conditions of a production plant that is specializing in manufacturing of the aluminum components for the automotive industry. The implementation framework of the DT for engine block manufacturing processes consists of three layers: physical layer, virtual layer and information-processing layer. A simulation model was created using the Tecnomatix Plant Simulation (TPS) software. In order to obtain real-time status data of the production line, programmable logic control (PLC) sensors were used for raw data acquisition. To increase production line productivity, the algorithm for bottlenecks detection was developed and implemented into the DT. Despite the fact that the implementation process is still under development and only partial results are presented in this paper, the DT seems to be a prospective real-time optimization tool for the industrial partner.

Keywords: digital twin; manufacturing; engine block

1. Introduction

The digital twin (DT) is becoming progressively more interesting to academia and industry. The main questions are: what technologies are needed to create DT and how can the DT improve manufacturing processes? Commonly characterized as containing a physical entity, a virtual counterpart, and the data connections in between, the DT is increasingly being examined as a way of advancing the efficiency of the physical assets. [1]. The DT is a digital description of an operating product or service that consist of its preferred features, attributes, circumstances, and ways of behaving through replicas, information, and data within a single or even across multiple life-cycle phases [2]. The DT has important pertinence in the production context as it has the capability to improve production processes, discover points of congestion in a production system, confirm settings, and simulate situations to forecast performance [3].

Manufacturing and fully integrated DTs are suggested in the literature. However, currently operated DTs in industrial practice still need attention [4]. According to Qi et al. [5] regardless of an intense ambition from small and medium-size enterprises (SMEs) to integrate DT into their everyday business, a majority of the SMEs are inexperienced with the essential technologies and tools of the DT. Feofanov and Baranov [6] formulated four obstacles to the implementation of the DT: lack of human resources, organizational barriers, high expenses, and the deficiency of legislative support. Furthermore, the next limitation relates to DT implementation occurring because there is no standardized approach for development of the DT, and an application framework for DT is missing. Normalized procedures guarantee comprehension between users while ensuring information flow between each stage of DT development and implementation [4]. However, the four standards for a DT manufacturing network [7–10] are currently under development. These standards define general principles of the DT framework for manufacturing and specify requirements for the reference architecture, digital representation of the manufacturing elements and information exchange.

Although many authors proposed and described various frameworks for DT implementation in the manufacturing industry context, there is an absence of real-life implementation studies reported in the available literature. The main aim of this paper is to demonstrate the feasibility of DT implementation under real conditions of a production plant specializing in manufacturing of the aluminum components for the automotive industry.

2. Literature Review

Even though the DT is considered as a challenging technology, it is still at the conceptual stage and only a few studies have specifically discussed methods for its development and implementation in the manufacturing domain [11–13]. Zheng et al. [12] proposed an application framework of the DT for welding production line. Zhang et al. [14] developed a DT-based approach for designing and multi-objective optimization of the hollow glass production line. Cao et al. [15] established a compressor manufacturing system architecture based on digital twinning. Bao et al. [16] proposed an approach to modelling and operations for the DT in the context of manufacturing. In order to provide the implementation methods of virtual-physical convergence and information integration for a factory, they used Automation Markup Language for modelling a structural parts machining cell. Liu et al. [17] developed a DT-driven methodology for rapid individualized design of the automated flow-shop manufacturing system. Ding et al. [18] presented the framework reference model of a DT-based Cyber-Physical Production System (DT-CPPS) and discussed in detail its configuring mechanism, operating mechanism, and real-time data-driven operations management. Zhou et al. [19] proposed a general framework for a knowledge-driven DT manufacturing cell towards intelligent manufacturing, which could support autonomous manufacturing by an intelligent perceiving, simulating, understanding, predicting, optimizing, and controlling strategy. Zhuang et al. [20] presented a detailed implementation process of the proposed framework of the DT-based smart production management and control approach in a satellite assembly shop-floor scenario. Bilberg and Malik [21] developed a DT-driven human-robot assembly system that extends the use of virtual simulation models developed in the design phase of a production system to the operations for real-time control, dynamic skill-based tasks allocation between human and robot, sequencing of tasks, and developing a robot's program accordingly. He and Bai [22] proposed the framework of DT-driven sustainable intelligent manufacturing. To provide guidance for practical development and implementation, Liu et al. [23] developed a conceptual framework for DT-enabled collaborative data management for metal additive manufacturing systems, where a cloud DT communicates with distributed edge DTs in different product life-cycle stages. Zamba et al. [24] proposed DT for the manufacturing of structural parts made from carbon fiber composite materials. In order to optimize production process, Sujová et al. [25] developed the DT of the real assembly line. Cimino et al. [26] proposed a practical implementation of the DT in manufacturing an execution system equipped within an assembly laboratory line.

3. Materials and Methods

Implementation framework of the DT for engine block manufacturing processes consists of three layers: physical layer, virtual layer and information-processing layer. The bidirectional mapping and interoperability of a physical space and virtual space are realized through data interaction [12].

3.1. Characteristics of Production Line

The process flow for aluminum-alloy engine block manufacturing includes melting, holding casting, fettling, heat treatment, and machining. However, the subject of the present study is only the machining part of the production line. Once the engine block is removed from the casting process a number of different machining processes need to be performed. The engine block machining line consists of a robotic arm, two horizontal machining centers, conveyor system, and buffer system (see Figure 1).

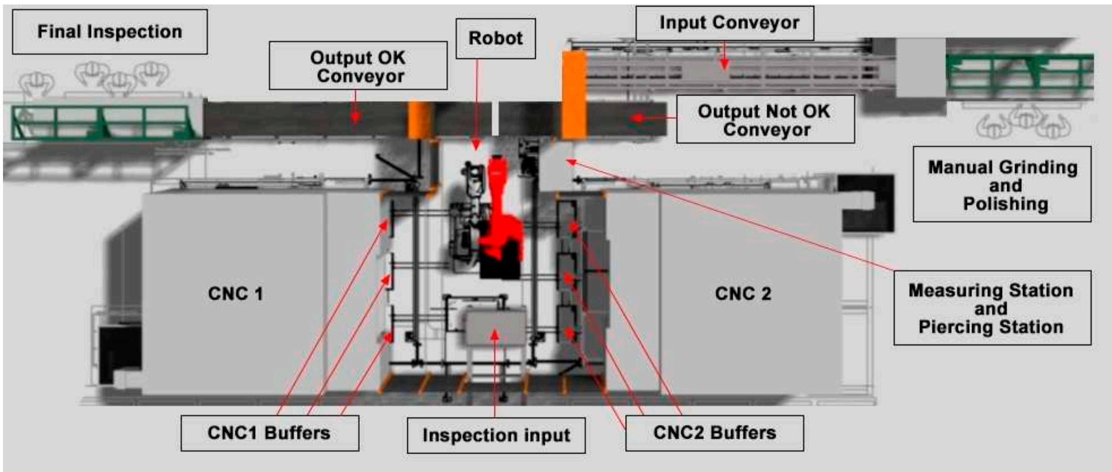


Figure 1. Machining cell layout.

After loading of the serial number of the casting by the line scan camera fixed on the input conveyor, the robotic arm is used to transport the workpiece from the conveyor to the empty buffer. The robot's decision-making algorithm is shown in Figure 2.

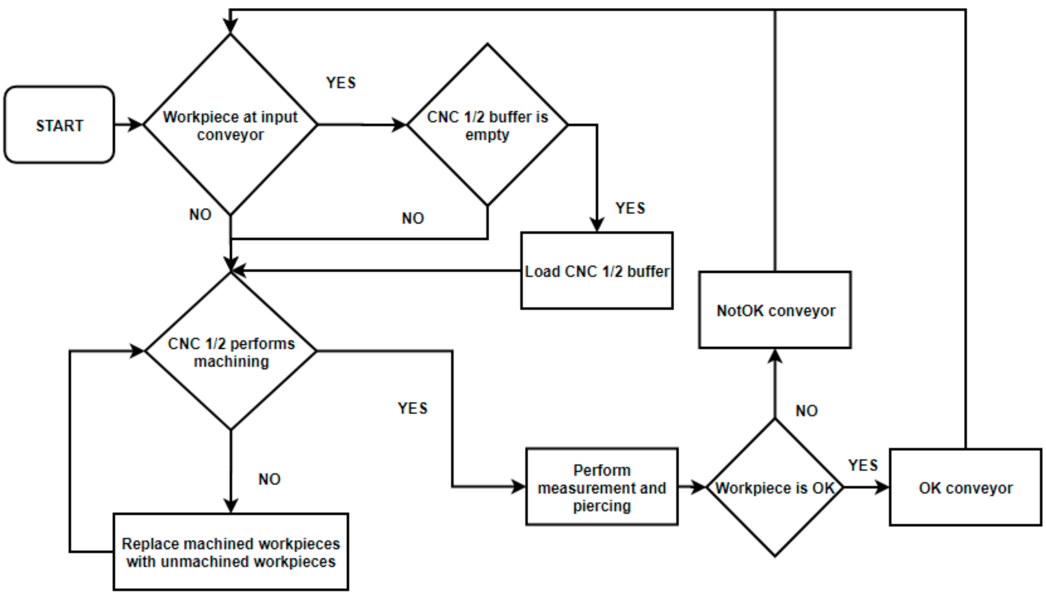


Figure 2. Robot's decision-making algorithm.

Engine block computer numeric control (CNC) machining includes processes such as milling, face milling, drilling, boring, tapping, threading, etc. Two clamping fixtures are positioned in the machining area of the CNC machine by a rotary table. Machining is followed by measurement and piercing of excess workpiece parts. After a trouble-free execution of all operations, the engine block is placed on the OK output conveyor. In case of an error, the workpiece is defined as defective. Therefore, manual processing is required and workpiece is placed on the Not OK conveyor.

3.2. Development of Simulation Model

Based on the aforementioned process flow, a simulation model was created using the Tecnomatix Plant Simulation (TPS) software (version 15, license TN75010-1253850, Siemens, Munich, Germany) (see Figure 3). The TPS software uses an object-oriented modeling method. Based on the library of typical equipment provided by the TPS, we customized the equipment that matches the actual situation. Control of the simulation process was achieved by programming with the SimTalk programming language.

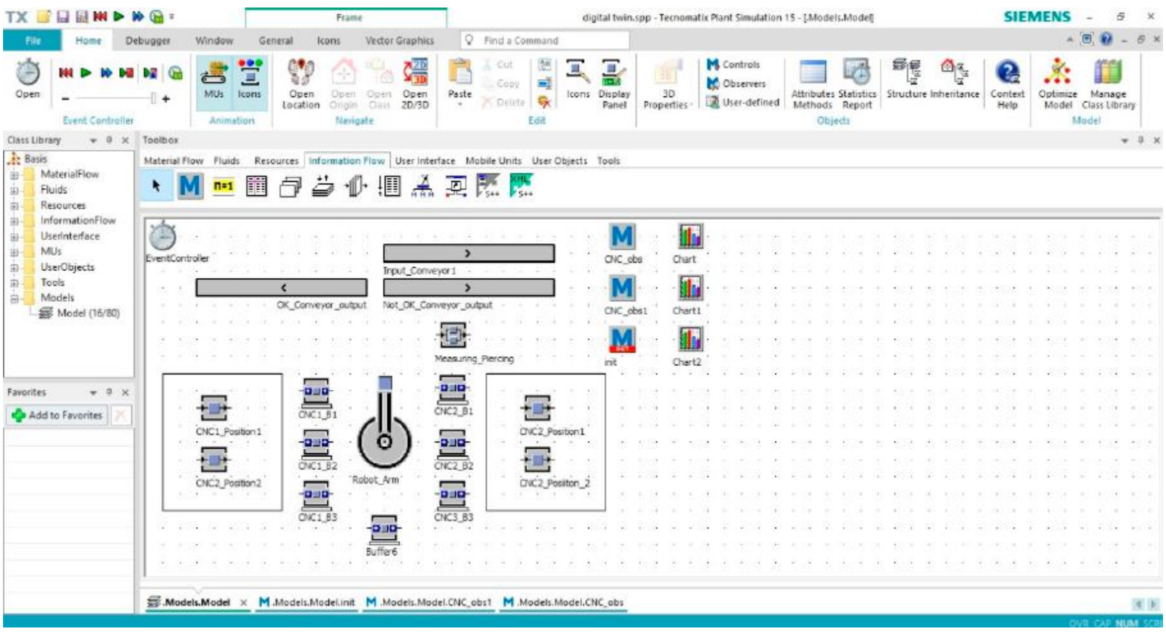


Figure 3. Development of simulation model in the Tecnomatix Plant Simulation environment.

Modeling of production systems in the TPS is realized by implementation of virtual objects, which represent individual production equipment, machines, workers, transport and logistics systems such as conveyors, trucks, loaders, etc., warehouses, buffers and other storage objects occurring in manufacturing companies. Hierarchically, the model of a production line or production system is created by sequential modeling in the TPS software that represents the states occurring on a real production line in a virtual world. The simulation model consists of active and static objects. Active objects include facilities that perform production or logistics operations. In our case, the Source, SingleProc, Buffer, PickandPlace (robot) and Conveyor objects are included among the active objects. The behavior and logical sequences of an operations are created using source codes, which are written in the Method interface. The individual methods are then assigned to the individual devices.

3.3. Development of Communication Interface

Data acquisition represents a crucial part of implementing the DT of engine block manufacturing processes. In order to obtain real-time status data of the production line, programmable logic control (PLC) sensors were used for raw data acquisition. Data were gathered from the distributed PLCs

(18 collection points) according to the configurable trigger conditions by data logger (Softing Industrial Automation GmbH, echocollect e, Haar, Germany). Verification of the PLC signals was performed by a software system (Autem GmbH, PLC-Analyzer pro 6, Emden, Germany). Subsequently, collected data were transferred into the local Structured Query Language (SQL) database. Two types of network protocols were used: Transmission Control Protocol/Internet Protocol (TCP/IP) as transport protocol and Siemens S7 as application protocol. TPS has an Open Database Connectivity (ODBC) interface that is able to retrieve data of events from the physical layer via the cloud-based database. Data exchange between the SQL database and the TPS software was allowed using the object SQlite. Internal cloud platform was used as a cloud-based information repository. Used real data are collected and then implemented into the production line DT. Collected data from the production line include: the position of the robot, the occupancy of buffers in front of the CNC and inside the CNC, the occurrence of workpieces on conveyors, the handling time of the robot, the machining time of the CNC, and the time from the placement of the workpiece at the input of the conveyor to its exit from the conveyor. Synchronization between the physical and virtual worlds is achieved using PLC controllers, which are analyzed and tested by the external test software. DT mapping diagram between layers is shown in Figure 4.

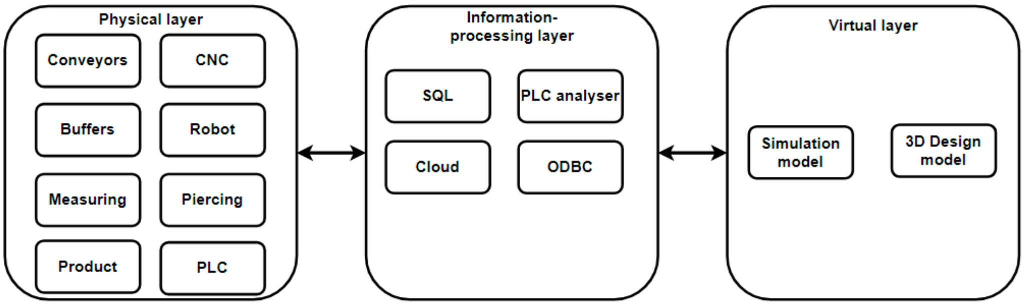


Figure 4. Digital twin mapping diagram between layers.

4. Results

Programmed source codes and methods to create a faithful replica of the physical production line are shown in Figure 5. To verify and validate conformity of the simulation model with the real production line, a test of the correct logical sequence of operations was performed without real production data firstly. The simulation model was compared with a video recording from the real production line to achieve compliance of the virtual twin with the real production line. An agreement between the real and virtual production line can also be observed when comparing real data from the production line with the Gantt chart generated by the DT. TPS software allows CAD (computer-aided design) models to be implemented in the production line interface in STEP (Standard for the Exchange of Product Data) or JT (Jupiter Tessellation) file format to enhance visual aspect of the DT. Detailed 3D simulation model of the production line in the TPS software is presented in Figure 6. This model is able to reproduce, in near real-time, all actions of the physical twin using feedback from the embedded sensors.

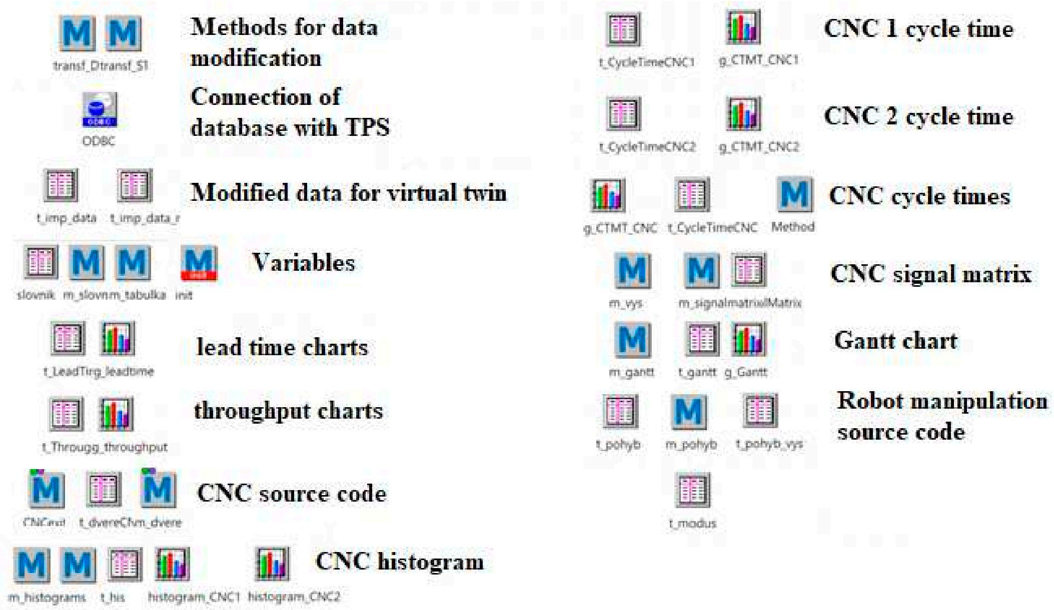


Figure 5. Source codes, methods and charts developed in the Tecnomatix Plant Simulation software.

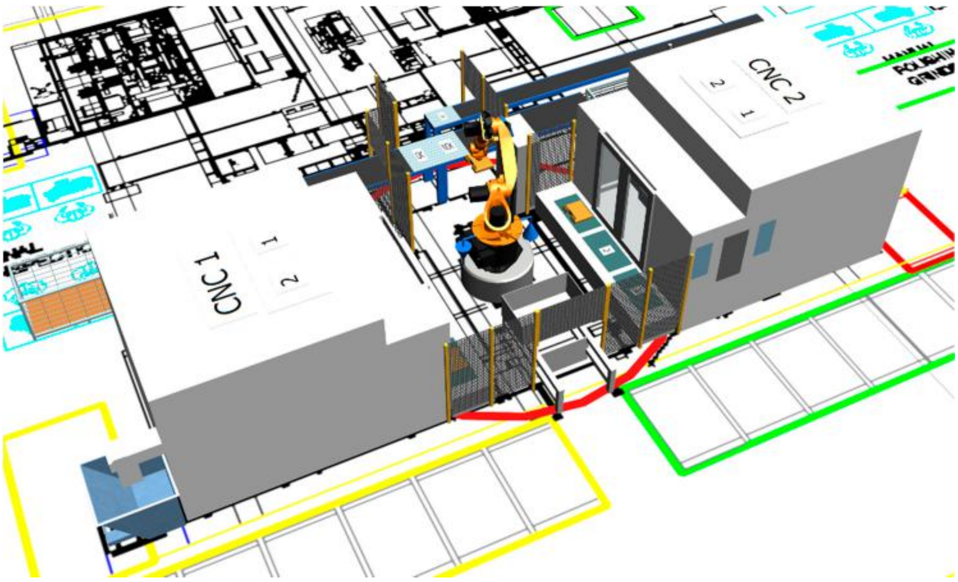


Figure 6. Machining cell visualisation in Tecnomatix Plant Simulation.

Data from the real production line, recorded using the Echocollect device, were implemented into the simulation model. Recorded data contained the following parameters: signal sequence number, signal occurrence time, manipulation time for each operation, robots position/process, data collection point description, connection of each signal to its corresponding device in the simulation model, and the total production time. Remote monitoring of the production line parameters was possible through visualization of:

- lead time (time interval from entry of the workpiece into the production process through the input conveyor to the removal of the workpiece from the output conveyor);
- cycle times and histograms for the CNC machines;
- throughput (the interval of occurrence of a casting at a given point);
- Gantt chart of the entire production line.

Remote monitoring through the DT refers to an ability to monitor and observer specific processes and operations within production without a physical presence in the production process. Remote monitoring involves observation of the physical production parameters through a dashboard or screen in a virtual world. Within DT, TPS software provides an interface for production line modeling with the ability to monitor physical production line status, while a personal computer represents viewing hardware.

To increase production line productivity, the algorithm for bottlenecks detection was developed and implemented into the DT. The DT uses real-time data from the production line to detect bottlenecks. Implementation and synchronization of the data between the physical and virtual worlds in the DT is in real time, compared to the classical simulation that uses data that are implemented offline and only historical data can be implemented. Bottlenecks within a production line significantly reduce throughput of the system, so quick and correct identification of bottleneck locations can lead to an increase of system throughput. The algorithm for bottlenecks detection and prediction is based on the real-time production data. Bottleneck detection and diagnosis is one of the main roles of the DT. Based on the implemented algorithm to identify bottlenecks of the production line, it was possible to determine the following bottlenecks shown in Table 1. A bottleneck arises when the workload is higher than the workplace can handle in terms of capacity. In general, any system (conveyor, buffer, CNC machine, measuring station, piercing station, or robot) can become a bottleneck. The bottleneck detection algorithm reveals deviations from the standard processing, machining, or handling time. The DT records data on such occurrence into the table. Based on the analysis of the results obtained from the DT, we were able to identify following conditions that affects optimal operation of the production line the most:

Table 1. Production line bottlenecks occurrence.

Occurrence	Bottleneck
15	extended measuring time
62	full output OK conveyor (robot must wait to free up space on the conveyor)
66	extended time since the casting was placed on the OK conveyor until it left the OK conveyor
10	extended time since the casting was placed on the Not OK conveyor until it left the Not OK conveyor
24	extended time for which the casting is ready to be removed from the inspection input

Based on the data analysis from the DT, we were able to detect bottlenecks in the production line. The DT provides a new perspective on the production line behavior and its management. The original estimate before the introduction of the DT was that a main bottleneck is caused by the CNC machine itself, but the results from DT show that the greatest impact on the extension of the throughput of the production line are the output conveyors and the measuring station. In addition, output conveyors are influenced by human factors. The worker must remove the casting from the process at the end of the conveyor to make room for the robot to place the next casting. The worker is also in charge of controlling the conveyor. Late start-up of the conveyor prolongs the time of casting on the conveyor, which in turn causes occupancy and the robot must wait to free up space in order to continue performing the required operations.

5. Discussion

DT implementation is an important vision for today's shop floors, especially for SMEs who want to reach the goals of smart manufacturing by using minimum investments and manpower [11].

A limiting factor for the data collection was discovered during the DT's evaluation. Several data collection point signals in the virtual layer showed latencies in the connection. Latencies caused writing of faulty data properties, which resulted in missing motions in the database. Therefore, each data collection signal time impulse had to be modified to ensure correct data collection. A similar problem with latencies was reported by Redelinghuys et al. [27].

The DT provides the capability to simulate the production line in near-real time. The physical state of the production line can be acquired through the PLC sensors and the status of each part of the production line in form of raw data is stored in the Local SQL database and in the Cloud database. With the use of TPS software, the physical production line was simulated. Information-processing layer established communication through the ODBC interface between the physical and virtual worlds. The throughput diagram, histograms, cycle time chart, lead time diagram and Gantt chart are used for the remote monitoring of the physical state of the production line. A bottleneck detection algorithm was implemented in the DT, which provides a new insight into production line behavior.

6. Conclusions and Future Work

This paper demonstrated the feasibility of DT implementation in the real condition of a production plant that is specializing in the manufacturing of an aluminum components for the automotive industry. The main contributions of this paper can be summarized as follows:

1. The simulation model of the engine block machining process was developed and validated.
2. Real-time interaction between physical and virtual entities of the production line was established.
3. To increase production line productivity, an algorithm for the bottleneck detection was developed and implemented in the DT.

Despite the fact that the process of the DT implementation is still under development and only partial results were presented in this paper, the DT seems to be a prospective real-time optimization tool for an industrial partner. Several authors were able to create a DT, but mainly laboratory conditions are presented in literature. Sun et al. [28] was able to improve the throughput of a production line with development of the throughput prediction DT model to predict the future throughput with simplicity and high accuracy. Liau et al. [29] created a DT framework for injection-molding processes that helped in the real-time to optimize and monitor processes and predict defects and quality of the final product. Talkhestani et al. [30] created a DT with the ability to detect changes in the production system to reduce the occurrence of an errors in production.

The novelty of our approach is in implementation of the DT in practical scenario, compared to other authors who implemented DT only in laboratory conditions. Furthermore, the DT uses real-time data synchronization between virtual and physical world, compared to the conventional simulation that uses only historical data.

Future work will be focused in two directions. Firstly, to increase production line productivity, optimization of the production line based on the detected bottlenecks from the DT will be undertaken. Secondly, an investigation to identify potential security risks need to be performed in order to develop reliable and effective cyber-security measures against possible forms of cyber-attack.

Author Contributions: Conceptualization, M.D. and M.Š.; methodology, M.D. and L.K.; simulation models development, R.B; writing—original draft preparation, R.B. and M.D.; writing—review and editing, L.K. and R.B.; visualization, R.B.; supervision, L.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jones, D.; Snider, C.; Nassehi, A.; Yon, J.; Hicks, B. Characterising the Digital Twin: A systematic literature review. *CIRP-JMSTP* **2020**, *29*, 36–52. [CrossRef]
2. Stark, R.; Damerau, T. Digital Twin. In *CIRP Encyclopedia of Production Engineering*; Chatti, S., Tolio, T., Eds.; The International Academy for Production Engineering; Springer: Berlin, Germany, 2019.

3. Roy, R.B.; Mishra, D.; Pal, S.K.; Chakravarty, T.; Panda, S.; Chandra, M.G.; Pal, A.; Misra, P.; Chakravarty, D.; Misra, S. Digital Twin: Current scenario and case study on manufacturing process. *Int. J. Adv. Manuf. Technol.* **2020**, *107*, 3691–3714. [CrossRef]
4. Fuller, A.; Fan, Z.; Day, C.; Barlow, C. Digital Twin: Enabling Technologies, Challenges and Open Research. *IEEE Access* **2020**, 108952–108971. [CrossRef]
5. Qi, Q.; Tao, F.; Hu, T.; Anwer, N.; Liu, A.; Wei, Y.; Wang, L.; Nee, A.Y.C. Enabling technologies and tools for digital twin. *J. Manuf. Syst.* **2020**. [CrossRef]
6. Feofanov, A.; Baranov, N. Risk analysis in digital twin creation of machine building production. In *MATEC Web of Conferences*; EDP Sciences: Les Ulis, France, 2019; Volume 298.
7. ISO/DIS 23247-1. *Automation Systems and Integration. Digital Twin Framework for Manufacturing. Part 1: Overview and General Principles*; International Organization for Standardization: Geneva, Switzerland, 2020.
8. ISO/DIS 23247-2. *Automation Systems and Integration. Digital Twin Framework for Manufacturing. Part 2: Reference Architecture*; International Organization for Standardization: Geneva, Switzerland, 2020.
9. ISO/DIS 23247-3. *Automation Systems and Integration. Digital Twin Framework for Manufacturing. Part 3: Digital Representation of Manufacturing Elements*; International Organization for Standardization: Geneva, Switzerland, 2020.
10. ISO/DIS 23247-4. *Automation Systems and Integration. Digital Twin Framework for Manufacturing. Part 4: Information Exchange*; International Organization for Standardization: Geneva, Switzerland, 2020.
11. Tan, Y.; Yang, W.; Yoshida, K.; Takakuwa, S. Application of IoT-Aided Simulation to Manufacturing Systems in Cyber-Physical System. *Machines* **2019**, *7*, 2. [CrossRef]
12. Zheng, Y.; Yang, S.; Cheng, H. An application framework of digital twin and its case study. *J. Ambient. Intell. Humaniz. Comput.* **2019**, *10*, 1141–1153. [CrossRef]
13. Negri, E.; Fumagalli, L.; Macchi, M. A Review of the Roles of Digital Twin in CPS-based Production Systems. In *Value Based and Intelligent Asset Management*; Márquez, A.C., Macchi, M., Parlikad, A.K., Eds.; Springer: Cham, Switzerland, 2019; pp. 291–307.
14. Zhang, H.; Liu, Q.; Chen, X.; Zhang, D.; Leng, J. A Digital Twin-Based Approach for Designing and Multi-Objective Optimization of Hollow Glass Production Line. *IEEE Access* **2017**, *5*, 26901–26911. [CrossRef]
15. Cao, J.; Wang, J.; Lu, J. A Referenced Cyber Physical System for Compressor Manufacturing. *MATEC Web Conf.* **2020**, *306*, 02005. [CrossRef]
16. Bao, J.; Guo, D.; Li, J.; Zhang, J. The modelling and operations for the digital twin in the context of manufacturing. *Enterp. Inf. Syst.* **2019**, *13*, 534–556. [CrossRef]
17. Liu, Q.; Zhang, H.; Leng, J.; Chen, X. Digital twin-driven individualized designing of automated flow-shop manufacturing system. *Int. J. Prod. Res.* **2019**, *57*, 3903–3919. [CrossRef]
18. Ding, K.; Chan, F.T.S.; Zhang, X.; Zhou, G.; Zhang, F. Defining a Digital Twin-based Cyber-Physical Production System for autonomous manufacturing in smart shop floors. *Int. J. Prod. Res.* **2019**, *57*, 6315–6334. [CrossRef]
19. Zhou, G.; Zhang, C.; Li, Z.; Ding, K.; Wang, C. Knowledge-driven digital twin manufacturing cell towards intelligent manufacturing. *Int. J. Prod. Res.* **2020**, *58*, 1034–1051. [CrossRef]
20. Zhuang, C.; Liu, J.; Xiong, H. Digital twin-based smart production management and control framework for the complex product assembly shop-floor. *Int. J. Adv. Manuf. Technol.* **2018**, *96*, 1149–1163. [CrossRef]
21. Bilberg, A.; Malik, A.A. Digital twin driven human-robot collaborative assembly. *CIRP Ann.* **2019**, *68*, 499–502. [CrossRef]
22. He, B.; Bai, K.J. Digital twin-based sustainable intelligent manufacturing: A review. *Adv. Manuf.* **2020**, 1–21. [CrossRef]
23. Liu, C.; Le Roux, L.; Körner, C.; Tabaste, O.; Lacan, F.; Bigot, S. Digital Twin-enabled Collaborative Data Management for Metal Additive Manufacturing Systems. *J. Manuf. Syst.* **2020**. [CrossRef]
24. Zambal, S.; Eitzinger, C.; Clarke, M.; Klintworth, J.; Mechin, P. A digital twin for composite parts manufacturing: Effects of defects analysis based on manufacturing data. In Proceedings of the IEEE 16th International Conference on Industrial Informatics (INDIN), Porto, Portugal, 18–20 July 2018; pp. 803–808.
25. Sujová, E.; Čierna, H.; Zabinska, I. Application of digitization procedures of production in practice. *Manag. Syst. Prod. Eng.* **2019**, *27*, 23–28. [CrossRef]
26. Cimino, C.; Negri, E.; Fumagalli, L. Review of digital twin applications in manufacturing. *Comp. Ind.* **2019**, *113*, 103130. [CrossRef]

27. Redelinghuys, A.J.H.; Basson, A.H.; Kruger, H. A six-layer architecture for the digital twin: A manufacturing case study implementation. *J. Intell. Manuf.* **2020**, *31*, 1383–1402. [CrossRef]
28. Sun, H.; Li, C.; Fang, X.; Gu, H. Optimized throughput improvement of assembly flow line with digital twin online analytics. In Proceedings of the 2017 IEEE International Conference on Robotics and Biomimetics (ROBIO), Macau, China, 5–8 December 2017; pp. 1833–1837.
29. Liau, Y.; Lee, H.; Ryu, K. Digital twin concept for smart injection molding. *IOP Conf. Ser. Mater. Sci. Eng.* **2018**, *324*, 012077. [CrossRef]
30. Talkhestani, B.A.; Jazdi, N.; Schlogl, W.; Weyrich, M. A concept in synchronization of virtual production system with real factory based on anchor-point method. *Proceedia CIRP* **2018**, *67*, 13–17. [CrossRef]

applied sciences

MDPI

Article

Towards Semi-Automatic Generation of a Steady State Digital Twin of a Brownfield Process Plant

Seppo Sierla [1], Lotta Sorsamäki [2], Mohammad Azangoo [1,*], Antti Villberg [3], Eemeli Hytönen [2] and Valeriy Vyatkin [1,4]

1. Department of Electrical Engineering and Automation, Aalto University, 00076 Aalto, Finland; Seppo.Sierla@aalto.fi (S.S.); Valeriy.Vyatkin@aalto.fi (V.V.)
2. VTT Technical Research Centre of Finland Ltd., 02044 Espoo, Finland; Lotta.Sorsamaki@vtt.fi (L.S.); Eemeli.Hytonen@vtt.fi (E.H.)
3. Semantum Oy, 02150 Espoo, Finland; Antti.villberg@semantum.fi
4. Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, 97187 Luleå, Sweden; valeriy.vyatkin@ltu.se

* Correspondence: Mohammad.azangoo@aalto.fi; Tel.: +358-465-819-664

Received: 26 August 2020; Accepted: 1 October 2020; Published: 5 October 2020

check for updates

Featured Application: A laboratory water heating and pressurizing process is used as a case study to demonstrate the proposed methodology for digital twin generation.

Abstract: Researchers have proposed various models for assessing design alternatives for process plant retrofits. Due to the considerable engineering effort involved, no such models exist for the great majority of brownfield process plants, which have been in operation for years or decades. This article proposes a semi-automatic methodology for generating a digital twin of a brownfield plant. The methodology consists of: (1) extracting information from piping and instrumentation diagrams, (2) converting the information to a graph format, (3) applying graph algorithms to preprocess the graph, (4) generating a simulation model from the graph, (5) performing manual expert editing of the generated model, (6) configuring the calculations done by simulation model elements and (7) parameterizing the simulation model according to recent process measurements in order to obtain a digital twin. Since previous work exists for steps (1–2), this article focuses on defining the methodology for (3–5) and demonstrating it on a laboratory process. A discussion is provided for (6–7). The result of the case study was that only few manual edits needed to be made to the automatically generated simulation model. The paper is concluded with an assessment of open issues and topics of further research for this 7-step methodology.

Keywords: digital twin; industrial process; steady state simulation; directed graph; piping and instrumentation diagram; Balas®

1. Introduction

Industrial process plants in sectors such as oil & gas, chemical, pulp & paper, power & heat, mineral processing, and water supply management have lifecycles of several decades. Retrofits offer a large potential for reductions in operating costs [1], energy consumption [2], CO_2 emissions [3], freshwater consumption [4], and environmental pollution [5]. The said authors proposed various kinds of models for assessing these reductions at design phase. However, no such models exist for the great majority of brownfield process plants, due to the considerable engineering effort involved [6]. In this article a *brownfield* is defined as an operating plant, which has existing physical structures and legacy software systems. The plant design information at a brownfield plant is generally not in digital format [7].

A *digital twin* is a special kind of plant model that has been synchronized with the physical process using recent sensor information. Thus, a digital twin would be especially suited for designing retrofits for brownfield plants that have been in operation for a long time. The following requirements are identified for the digital twin:

1. It accurately captures aspects of the plant relevant to the retrofit
2. It can be generated from source information commonly available at brownfield plants
3. Minimal manual engineering effort should be involved

Despite much recent research on digital twins, there is a lack of research addressing these requirements. Numerous definitions for digital twins have been proposed. An *experimentable* digital twin, based on a simulation model of the plant, is suited for assessing impacts of a retrofit [8]. Simulation approaches for industrial process plants can be categorized into *steady state* and *dynamic*. Dynamic simulation has the special capability of determining how the process state changes over time in response to an event such as the closing of a valve or a setpoint change. Such capabilities are essential for investigating modernization of automation systems. A dynamic simulation model can be extended to an experimentable digital twin through tracking approaches that synchronize the simulation model with the process state measured by sensors [9]. Essential source information for generating a high-fidelity dynamic simulation model includes pipeline routing details, which are used to determine pressure head losses [10]. If such information is available, for example from 3D CAD models, a dynamic simulation model can be automatically generated [10], and it can be extended to an experimentable digital twin with tracking simulation [11].

However, such source information is generally not available for brownfield plants [7,12]. Point clouds from 3D scanning of industrial plants support use cases such as detection of whether a factory layout is collision free [13], but this does not capture essential information for creating a dynamic simulation model, namely individual components and their connections. Thus, the focus of this article is experimentable digital twins based on steady state simulation models. Such twins could be used for supporting the operators in their daily decision-making or the management in strategic decision-making [14]. These twins may also be used for "what if"-studies, i.e., for the assessment of retrofits involving physical process configuration changes (e.g., process stream re-arrangements; removing, replacing or installing a new process equipment such as a purification step or heat recovery system) or changes in the key process parameters (e.g., temperature, consistency). As an outcome, the digital twin would evaluate the impact of design alternatives for the retrofit in terms of the process's fresh water, energy, chemical or utility consumption. It could also be used to determine the chemical state of the process by modelling the pH, COD levels (Chemical Oxygen Demand), TSS levels (Total Suspended Solids) or trace component amounts in the process streams. A steady state digital twin would be a powerful tool to improve understanding of the process, investigate abnormal situations in the plant or train process operators [14].

This article is structured as follows. Section 2 reviews related research in the fields of steady state simulation, digital twin research in the context of brownfield plants, and automatic generation of digital twins. Section 3 presents an overview of a methodology for generating steady state digital twins, and positions prior research and the contribution of this article in the context of the overview methodology. Further, Section 3 details the contribution in general terms as an object-oriented design. Section 4 applies the proposed methodology to a case study, a laboratory process. Section 5 summarizes the results as the key findings from the case study. Section 6 discusses the generalizability of the findings for other case studies, other steady state simulation tools, and to plants with varying degrees of digitally available engineering design information. Section 7 concludes the paper and identifies topics for further research.

2. Literature Review

2.1. Steady State Simulation

Steady state simulation is based on first principles such as conservation laws, phase equilibria, heat and mass transfer, and reaction kinetics. Steady state simulation focuses on stable operating conditions. Unlike dynamic simulation, it does not consider the time dependency of the process [14,15], so it assumes that variables are constant with respect to time. In steady state, there is no accumulation of mass or energy within the system, so the overall mass and energy input equals its output. Steady state simulation is typically conducted in the early-state design of plant wide systems or process departments. Steady state modeling and simulation has been widely used in the industry for establishing mass and energy balances, evaluating and improving the process performance, process design, plant equipment sizing, and process optimization [14–18]. Inputs to steady state models are pressures, temperatures, flows, and compositions; outputs are equipment sizing and process optimizations [14].

The computational complexity increases considerably from steady state to dynamic simulation. Thus, dynamic simulation model building requires significant additional engineering effort to determine the model parameters [18,19]. However, since the steady state simulation model is a basis for the development of a dynamic model [18], dynamic modelling can be considered only later when more understanding of the intended commercial implementation of the technology is available.

The level of detail in steady state simulation studies varies from small-scale chemical reactions to mill-wide process calculations in many fields of industry. In the pulp and paper industry, steady state simulation has been applied to optimize water consumption [20,21], minimize energy and utility consumption [22–24], and evaluate the chemical state of the process [25,26]. Kangas et al. [27] defined a steady state simulation model of a kraft pulp mill and evaluated the economic feasibility of the process. In the field of biorefineries, steady state simulation has been used for process modelling and evaluating the economic performance of biomethane [28], bioethanol [29], biodiesel [30], and renewable diesel production [31]. Barbosa et al. [32] used steady state simulation to study carbon capture and utilization opportunities in a sugarcane biorefinery. Hytönen and Stuart [33] used plant-wide steady state process simulation models as part of a methodology for early stage screening of forest biorefinery retrofit scenarios. Steady state simulation has also been used in numerous other process and/or economic performance studies in the field of wastewater purification [34], chemical production [18,35,36], mineral processing [37] and food industry [38,39].

2.2. Digital Twins for Brownfield Process Plants

A digital twin is an online replica of a physical system. Twins generally have a capability for synchronization with current sensor values of the system and in some cases the twin may impact the physical system through actuation [40]. Most of the research on digital plants considers greenfield sites with extensive information available in a digital format (e.g., [41–44]). However, the advantages of the digital twins are not limited to the greenfield plant. They can also improve brownfield plants economically, politically and environmentally [45]. The research on digital twins for brownfield process plants is limited and scattered, focusing on diverse topics such as evolving a manufacturing system with changing product requirements [46], determining whether a layout is collision free [13], upgrading control and data acquisition systems to Industrial Internet [47] and extracting knowledge from legacy documentation of industrial plants [12,48].

Sorensen et al. [46] present a digital platform of a brownfield manufacturing system which can handle changing product requirements. Shellshear et al. [13] use point cloud information obtained from 3D scanning of the factory floor to update information about collision free spaces. The results obtained in [49] show the benefits of using data driven approaches to generate a self-aware digital twin for process plants. It presents a method based on data-driven modelling that performs Big data analytics on process history data to improve process control efficiency. Makarov et al. [50] introduce a three step modelling process for a manufacturing system digital twin: the development of

SysML (Systems Modeling Language) diagrams, using AnyLogic as a tool for simulation modeling, and communicating with actual systems through the MES (Manufacturing Execution System). In a recent paper by Kychkin et al. [51], a method for digital twin implementation based on estimation of simulation model parameters and calculation of control signals for a dynamic ventilation system of underground mines was discussed. By considering the dynamics of air distribution and changes in environmental parameters, the proposed algorithms can improve safety and energy saving in mines, in which the ventilation process consumes from 30 to 50 percent of all company electricity.

2.3. Automatic Generation of Digital Twins

The research on automatic generation of digital twins has been motivated by several use cases. The closest state-of-the-art works and their differences are analyzed as follows. A dynamic digital twin, as defined in Section 1, has been generated from 3D CAD information for the purpose of using process state values from the twin as soft sensors [11]; the approach is not applicable to brownfield plants for which 3D CAD models are generally not available. A qualitative digital twin of the plant has been generated for co-simulating control software against the plant in order to detect logic errors in the virtual commissioning phase [12,48]. A digital twin has been generated for hardware- in- the- loop testing of control software [52], another activity that is not applicable in the context of steady state models that does not capture time-related behavior. A digital twin has been generated for the analysis of bottlenecks [53]; out of all the works reviewed in this section, this is the only one that is relevant to the specific purpose of the research presented in this paper, which is the design and validation of retrofits. However, the bottleneck analysis was performed specifically in the context of discrete manufacturing systems, so the approach is not applicable to the continuous processes addressed in this paper. Our main contribution over the state-of-the-art is to address the lack of research on automatic generation of digital twins, or even simulation models, that are applicable for the design and validation of retrofits to process plants.

Nowadays, a limited number of the most modern plants have digital, machine-readable design information available. The information for other plants is mostly accessible in printed papers, static PDFs, and other human-readable formats [7]. To limit the engineering cost of digital twin development for brownfield plants, it is important to have a fully or at least partially automatic solution for simulation model creation from the available plant information.

There are different available sources of information at process plants for the automatic generation of a digital display [11], such as datasheets, Process Flow Diagram (PFD) and Piping & Instrumentation Diagram (P&ID) diagrams, IO lists, 3D plant models and logic diagrams. The required information for simulation model creation can be extracted from these documents. The source information for the digital twin creation is not limited to design and engineering documents; for example, in [53] it was shown that a low fidelity digital twin has been generated automatically from high level requirements of the initial design phase of the project. Ref. [11] presents an automatic generation of simulation based digital twins for industrial process plants from 3D models. Sierla et al. [6] present an automatic solution to create the abstract graph model of the process system from a digital P&ID and a 3D CAD model of the system. This work was continued towards integrating the P&ID and CAD information by first converting the extracted information to the same level of abstraction [54]. Similarly, [55] introduces an automatic approach for matching 2D design documents and 3D scanned models of the process system by creating attribute graphs, calculating the level of similarity between graphs and merging the extracted results. Rantala et al. [56] use graph matching techniques to empower plant design engineers to reuse design information from existing process plants.

The digital twin is not the only use case for automatic model generation. For example, Son et al. [57] present a general automatic solution reconstruction of an as-built 3D model of a brownfield process plant from 3D laser-scan data, a 3D CAD database, and P&ID documents. An automatic solution for extracting information from laser-scan data to detect straight pipes, elbows, and tee pipes is presented in [58] for generating an as-built 3D pipeline model. However, there is a lack of work on using laser-scans as source information for experimentable digital twins. Further, it cannot be assumed that 3D CAD information is available for brownfield plants. Our use of steady state simulation models as the basis for experimentable digital twins simplifies the problem of obtaining source information: as is discussed in Section 3, a P&ID and access to recent process history can be sufficient source information for digital twins that are useful for the needs of retrofit projects.

3. Proposed Methodology

3.1. Methodology Overview

Figure 1 shows a vision for a methodology for the semi-automatic generation of a steady state digital twin for a brownfield process plant. The methodology consists of the following seven steps. Since several research works exist for steps 1 and 2, this article proposes solutions for the remaining steps.

1. The main process design document that is generally available at a brownfield process plant is a P&ID. Some leading P&ID CAD vendor's tools are able to export P&IDs in a machine readable format according to the standardized Proteus XML schema, but this capability is present only in the most recent tool versions and is thus not applicable to brownfield plants [6,59]. In general, the P&IDs at a brownfield plant are paper documents that have been scanned to pdf-format, so various image recognition techniques are needed to extract information from them [12,48,60,61].
2. Graphs have emerged as an intermediate format for abstracting key information from a process plant design [6,54–56]. The information that is relevant for building a steady state simulation model is extracted from the digitalized P&ID into a directed graph, as described in [6]. If Proteus XML is used as the digitalized P&ID format, the methodology will be able to support also modern plants for which the P&ID could be exported directly into this format. However, if the compatibility is not required, any proprietary format for a digitalized P&ID can be used as long as the graph is generated according to the following guidelines. Process equipment such as tanks, pumps, and valves are represented with nodes, and node labels capture the type of the component as well as the tag. Flows are represented with directed edges between the components, and the type of flow (e.g., water or broke) is captured by the edge label.
3. The graph should be transformed until it is at a level of abstraction in which the steady state model can be generated by performing a one-to-one mapping from the graph nodes and edges to the equipment and flows of a steady state model.
4. Simulation tool specific rules should be defined and implemented for generating a flowsheet of the steady state model automatically from the graph generated in step 3. The rules should be implemented by a custom software tool that writes its output into a format that can be imported to the selected simulation tool.
5. A steady state modelling expert should manually finalize the flowsheet of the generated steady state model, using his or her expert modelling knowledge that could not be formalized as rules in step 4.
6. A steady state modelling expert should manually initialize the steady state simulation model by defining the chemical components (i.e., water, pulp, air, steam, etc.) and selecting the calculation modules for each process equipment using his or her expert modelling knowledge that could not be formalized as rules in step 4. Further research could try to automate this step.
7. The selected calculation module defines the needed input values for parameterizing the process equipment. The steady state modelling expert performs the parameterization manually. Further research could try to automate this step. If the parameterization is performed according to recent sensor data from the process, the steady state model may be considered as a digital twin.

This article is scoped as follows. Previous research exists for steps 1 and 2 as has been cited above. This article will provide solutions for steps 3, 4, and 5 and apply them to a case study. Steps 6 and 7 will be addressed as further work in the final section of the article.

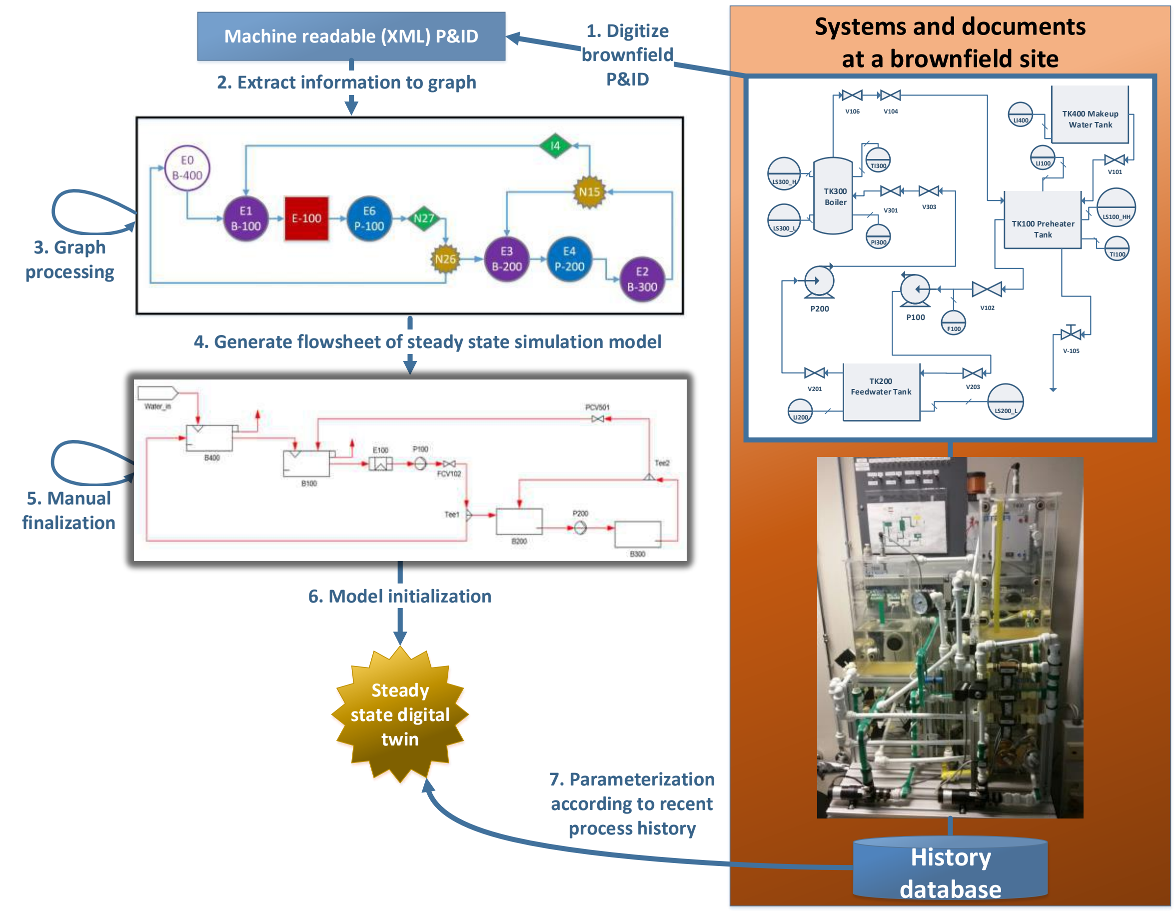


Figure 1. Proposed methodology for the semi-automatic generation of a steady state digital twin for a brownfield plant.

3.2. Graph processing

This section presents algorithms for step 3 of the methodology in Section 3.1. Figure 2 shows a Unified Modeling Language (UML) class diagram of the graph representation of the information extracted from the digitalized P&ID. This methodology assumes that such a graph has been previously generated, e.g., according to the approach presented in [6].

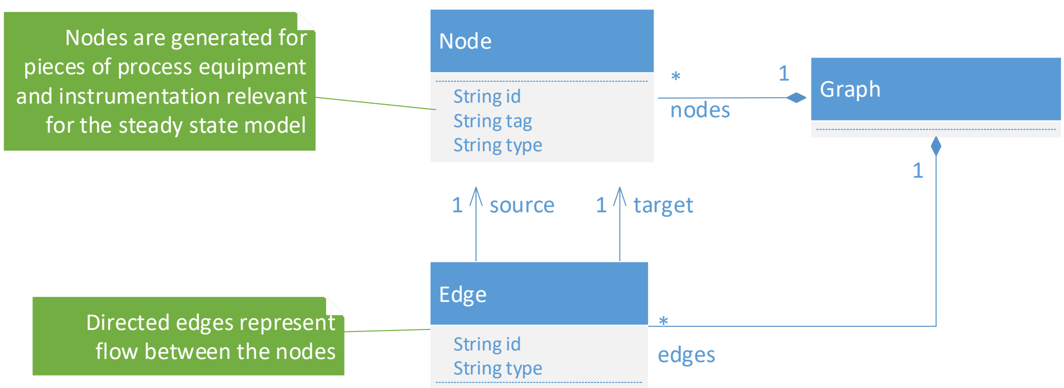


Figure 2. UML class diagram of the graph representation of the information extracted from the digitalized P&ID.

Dynamic modelling captures the time dependency of the process; it predicts all the transient states when the process moves from state A to state B [14]. Thus, the flowsheet of the dynamic simulation model is almost one-to-one with the P&ID including all the control loops with their control and binary valves. Steady state modelling, on the other hand, assumes that all variables are constant in spite of the ongoing process that tends to change them. Thus, the modelling of the binary valves of the control loops is irrelevant and it should not be included in the steady state simulation model. Control valves, on the other hand, are captured in the steady state simulation model to adjust the flows, temperatures, pressures, consistencies, etc. of the steady state. Figure 3, UML activity diagram represenation of an algorithm that removes binary valves from the intermediate graph representation of a process, shows an algorithm for removing the binary valves present in the P&ID from the intermediate graph presentation of the process. The algorithm iterates through all the nodes in the graph and uses node types to identify the binary valves. For each binary valve, the algorithm iterates through all of the edges in the graph to find the edges representing the outgoing and incoming flows of the valve, 'eDownstream' and 'eUpstream', respectively, in Figure 3. UML activity diagram represenation of an algorithm that removes binary valves from the intermediate graph representation of a process. The node and 'eDownstream' are removed. The target of 'eUpstream' is changed to the target of 'eDownstream'.

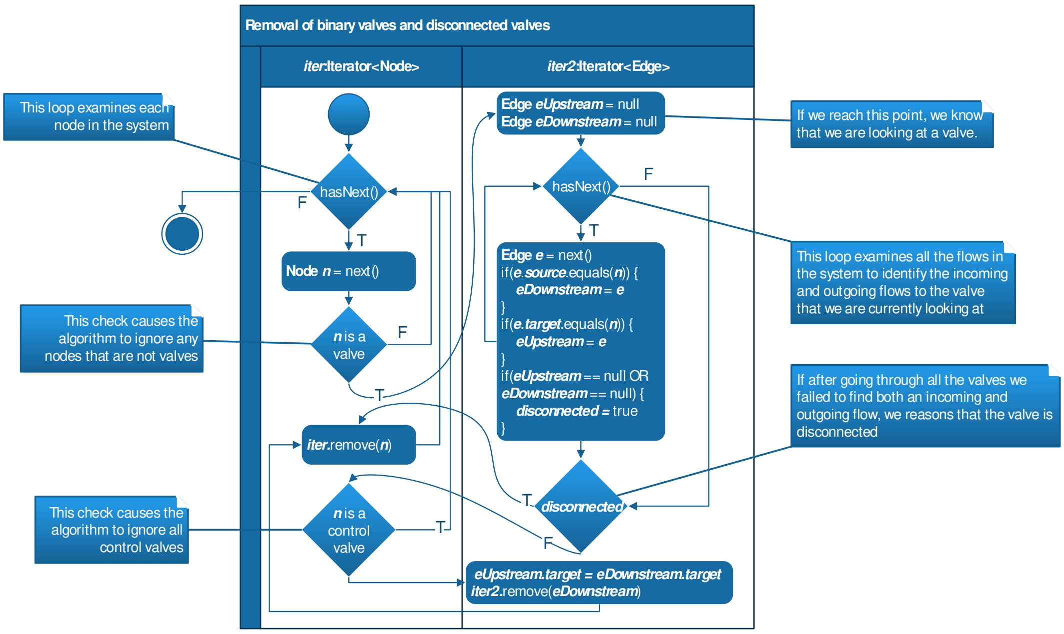


Figure 3. UML activity diagram represenation of an algorithm that removes binary valves from the intermediate graph representation of a process.

The binary valve removal algorithm in Figure 3 is relevant for all steady state simulation tools. In contrast, there are differences between how tools model tanks with internal heaters. In some tools, such as the tool used in our case study, the library of equipment symbols does not have a tank with an internal heater. In such cases, the tank with an internal heater can be modelled by adding the heater to the outgoing flow of the tank. Figure 4 shows an activity diagram for this purpose. The source information for this algorithm is the graph outputted by the algorithm in Figure 3, UML activity diagram represenation of an algorithm that removes binary valves from the intermediate graph representation of a process. The algorithm iterates through all nodes in the graph and looks for tanks with heaters ('heater' in Figure 4). For all such nodes, the algorithm iterates through all of the edges to find the outgoing edge. A node representing a heating element is added to the outgoing flow.

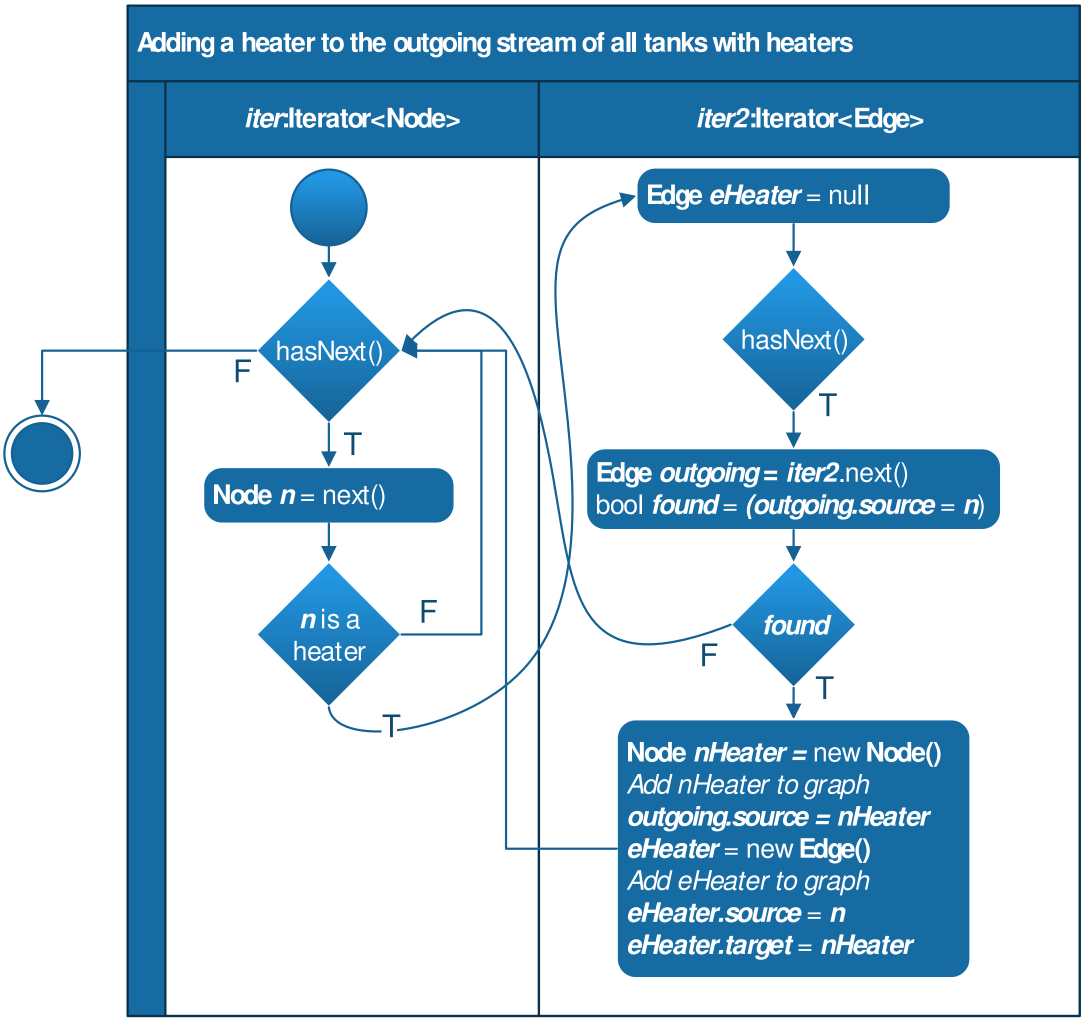


Figure 4. A UML activity diagram representation of an algorithm for manipulating heating elements in the intermediate graph.

3.3. Generating a Flowsheet of the Steady State Model

This section presents rules for step 4 of the methodology in Section 3.1. The rules are applied to the intermediate graph outputted by the algorithm in Figure 4. The rules (Table 1) are valid for the selected steady state simulation tool, Balas® (https://info.vttresearch.com/balas). Balas® is a steady state simulation package for chemical processes with emphasis on pulp and paper, food processing and biochemical processes. If another simulation tool was selected, a new set of rules should be created to correspond with the symbols of that simulation tool. Figure 5 presents some symbols in Balas® that are used to simulate different process equipment. Each symbol has one to several ports that are connected to either inlet or outlet streams. The implementation of the rules must ensure that each port of the symbol in the simulation tool is used at most once.

The rules are realized according to the object-oriented paradigm. Figure 6 shows a UML class diagram to capture these structures. Italics in the class diagram denote abstract classes and methods (i.e., the 'Component' class and its methods), so all inheriting classes (i.e., 'Pump', 'Heater', 'Splitter', 'Tank' and 'Valve') must implement these methods. The implementation of these methods should ensure that the ports are assigned according to the rules in Table 1 and that each port is used at most once.

Table 1. Rules for a one-to-one mapping from an intermediate graph to a steady state model.

Graph Structure	Mapping to Balas®
A node of type tank, with one outgoing edge and one or more incoming edges	Replace all tank nodes with a symbol "MDTank#1" (See Figure 5a). Add a stream from port#1 of symbol "Terminal in" (See Figure 5b) to port #2 of symbol "MDTank#1". From port #1 of symbol "MDTank#1", add a stream to nowhere. Ports #3-#12 of the symbol "MDTank#1" can be used either for feed or outlet.
A node of type valve, pump or heater with one incoming and one outgoing edge	The relevant symbols are "Valve" (See Figure 5c), "Pump" (See Figure 5d), "Heater/cooler#1" (See Figure 5e). For each of these symbols, port #1 is for inlet and port #2 is for outlet.
A node of type tee with one incoming and two outgoing edges	Replace tees with symbol "Splitter" (See Figure 5f) with port #3 for inlet and ports #1 and #2 for outlet.

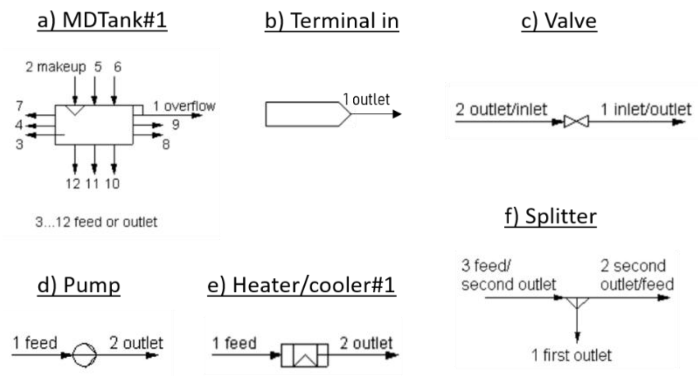


Figure 5. Different symbols in the Balas® simulation software for simulating process equipment.

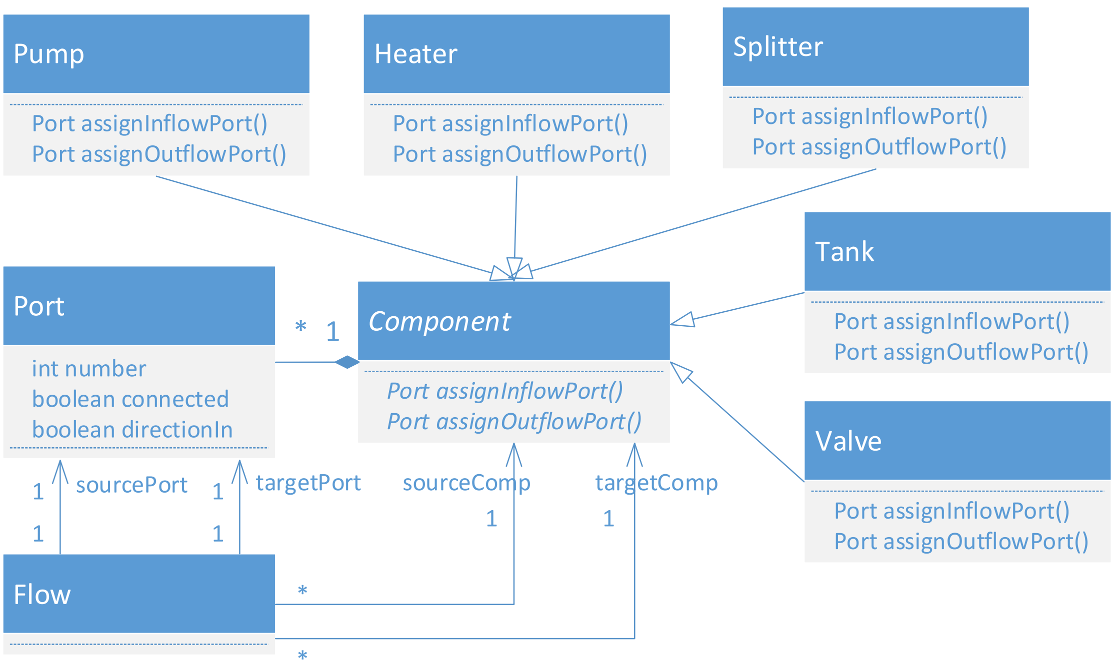


Figure 6. UML class diagram of the steady state model structure.

The rules in Table 1 are implemented by algorithms that map the intermediate graph outputted by the algorithm in Figure 4 to the object model in Figure 6. The creation of the various types of components is trivial as the type label of the node is used to determine whether to create an object of type 'Pump', 'Heater', 'Splitter', 'Tank' or 'Valve'. However, the connection of valves to the appropriate ports is not as straightforward. Figure 7 shows an algorithm for this purpose. The algorithm is general in the sense that changes are not required if more component types are added or if the port mapping rules are changed to meet the requirements of a specific steady state simulator. This generality was achieved by using abstraction and inheritance in Figure 6. Any changes are limited to the implementation of the methods 'assignInflowPort()' and 'assignOutflowPort()'.

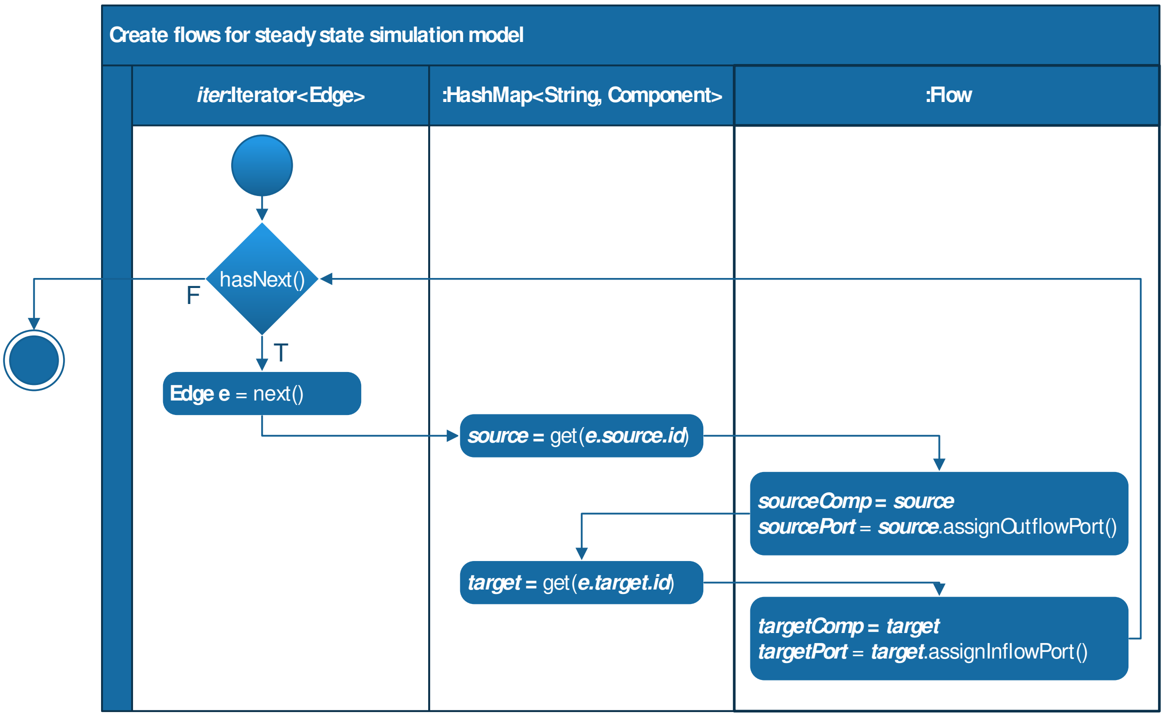


Figure 7. UML activity diagram representation of an algorithm that creates the flows and connects them to the correct ports according to the mapping rules in Table 1.

3.4. Implementation of the Design

The UML designs were implemented as follows in the Java programming language. The composition relation (line with a solid diamond ending) in the class diagrams was implemented with the Java *Vector* class. For example, the compositions in Figure 2 are implemented with a vector containing elements of type *Node* and *Edge*, i.e., *Vector<Node>* and *Vector<Edge>*. The *iterator()* method of these vectors are used to obtain the iterators *Iterator<Node>* and *Iterator<Edge>* in Figure 3. The triangular arrows in Figure 6 are generalization relationships. The algorithm in Figure 7 exploits the generalization, so that new component types, such as refiners, columns and reactors, can be added without changes to the algorithm.

4. Case Study

Aalto's water process plant, which is depicted in Figure 8, consists of different process, electrical, instrumentation and automation components. It can be used for demonstrating various process scenarios and related automation solutions for research and educational purposes. The main task of the plant is to supply heated and pressurized water for a variable load. The return water is reused in a closed-circuit stream. Five main closed control loops are defined to adjust the level of water in the tanks and the temperature and pressure of the supplied load.

The first process component of the closed loop primary stream of the process plant is a Preheater Tank (B-100), which receives the water that is returned from the Supplied Process. In the Preheater Tank, the temperature can be adjusted to the desired temperature by using a heater (E-100), a copper colored component in the bottom right tank in Figure 8. The Preheater Pump (P-100) transfers heated water from the Preheater Tank to the Feedwater Tank (B-200). From there, the Feedwater Pump (P-200) pressurizes the water in the Boiler (B-300) according to a setpoint value, despite disturbances caused by the Supplied Process. A makeup Stream compensates for the loss of water in the primary stream, which occurs gradually over time due to evaporation from the open tanks.

Figure 8. Aalto's water process plant.

The operator can run the plant in manual or automatic mode; to provide an automatic operation interface, all the sensors and actuators are connected to a remote I/O system, which transfers data between the plant and a soft PLC implemented on a PC. By using OPC UA, the field and automation data can be sent to simulation software like Balas®, Simulink, and Apros.

Figure 9 shows a P&ID of the case process. The P&ID was originally drawn in the SmartPlant P&ID tool but has been redrawn to reduce clutter. The P&ID was exported to Proteus XML using SmartPlant P&ID. The graph in Figure 10 was generated using the methodology in [6]. The graph was processed by the algorithms in Figures 3 and 4. The results are shown in Figures 11 and 12, respectively.

A flowsheet of the steady state model was generated from the graph in Figure 12 using the rules defined in Table 1 and the algorithm in Figure 7. The resulting object model confirming to the class diagram in Figure 6 was serialized to .csv format (Tables 2 and 3). It is visualized in Figure 13 to help the reader verify that the port numbers and connection conform to the rules in Table 1.

The .csv output was imported to the Balas® steady state simulation tool, resulting in the model in Figure 14. A custom importer plugin was created for a demonstration version of Balas® that is based on Simantics Open operating system for modelling and simulation (https://www.simantics.org/). It uses a graph database for storing simulation models and related data. Simantics provides a general-purpose functional scripting language SCL that is capable of manipulating the models within the database. SCL is also suitable for programming utility functionality on top of released simulation tool products. This version of Balas® includes an IDE for developing and testing SCL-based plugins within the simulator environment. Using SCL APIs for Balas®, a translator function was created that takes as input .csv files and creates corresponding Balas® model structures defined with flowsheet graphics. The importer plugin can in the future be extended to implement automation of steps 6–7 of the methodology.

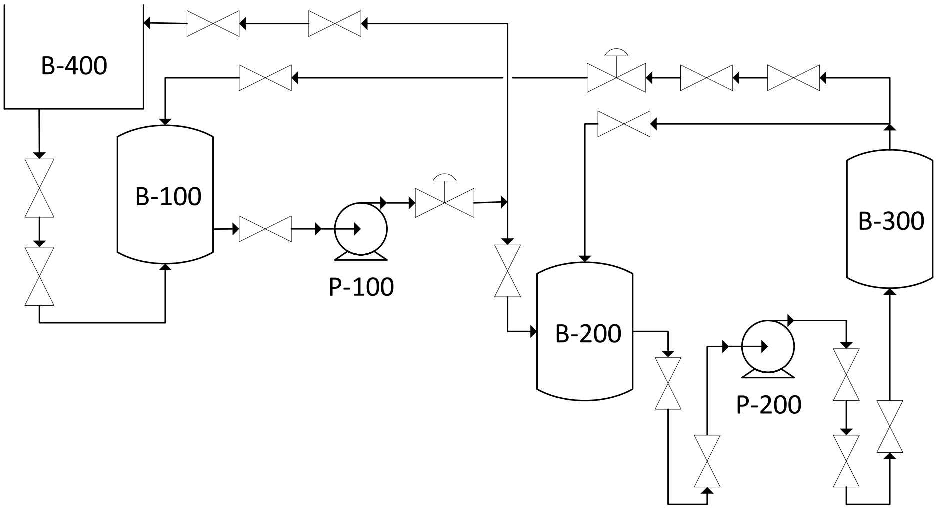


Figure 9. P&ID of the case process.

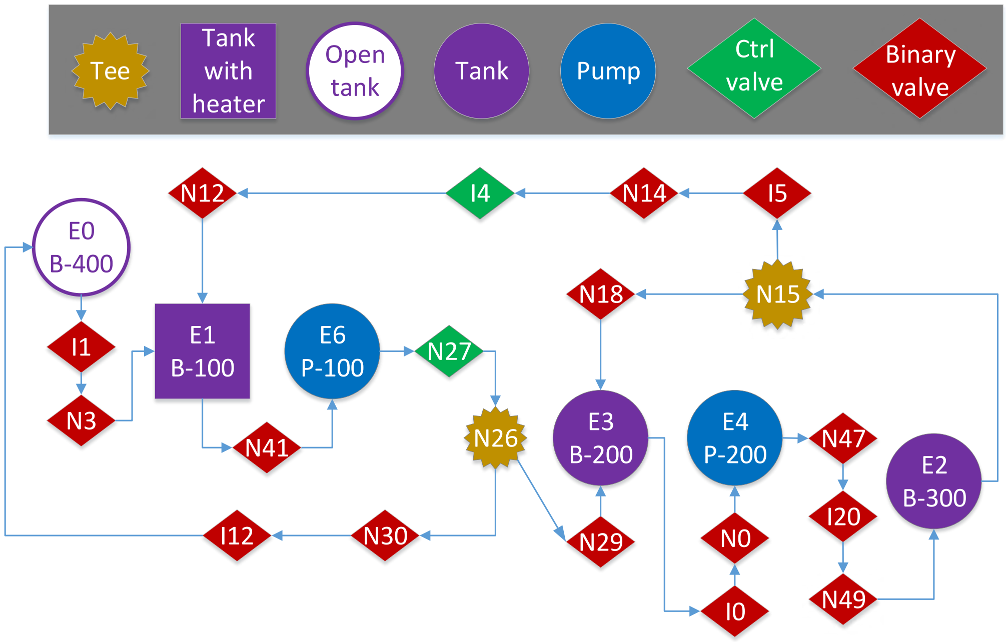


Figure 10. Graph representation of the information extracted from the digitalized P&ID.

Step 5 of the methodology in Section 3.1 involves a modeler making manual finalizations to the model in Figure 14 according to expert modelling knowledge that could not be captured as generally valid rules, such as the ones in Table 1. In the selected simulation tool, i.e., Balas®, there are two different kind of calculation modules available for simulating a normal tank.

The more complicated calculation module can be used to simulate a storage tank with several inflows and outflows, an overflow, and a makeup stream. This calculation module is used for simulating buffer tanks. During the simulated steady state, this buffer tank constantly provides a fixed outflow requested by the receiving module located after the buffer tank. If the required amount of flow is not available, the buffer tank provides the missing part through the makeup stream (port #2). The makeup stream may be connected to another tank or e.g., to the freshwater system. If the inflows of the buffer tank exceed the required outflow, the surplus is led to the overflow stream (port #1) which may be connected to another tank or alternatively to the drain. In the case example, the buffer tanks "B400" and "B100" are simulated using this calculation module. The case process is initially filled up with fresh tap water through the makeup stream ("Makeup1") of the makeup tank "B400". Also, if there are any leaks in the system, the makeup flow to cover the leaks is taken from the tap water line ("Water in"). The valve "FCV102" is the receiving module that requests a specific flow from the tank "B100". If the inflow to the tank "B100" from the tank "B300" through the valve "PCV501" is not sufficient, the tank "B100" requests makeup from the makeup tank "B400", and not from the fresh water system ("Makeup source 2"), as simulated in Figure 12. This change of the makeup stream source was done manually based on the expertise of the modeler.

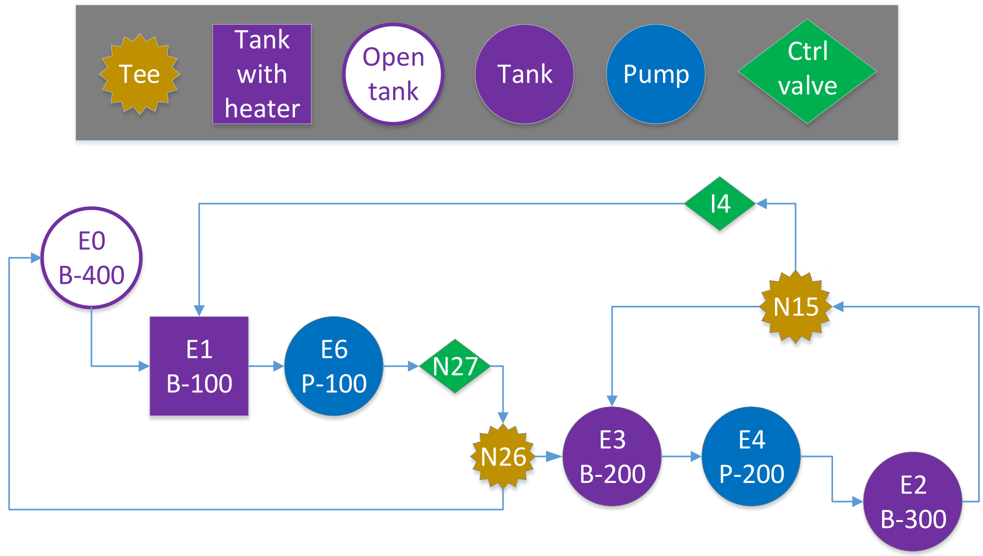


Figure 11. The result of processing the graph in Figure 10 with the algorithm in Figure 3. All binary valves are removed.

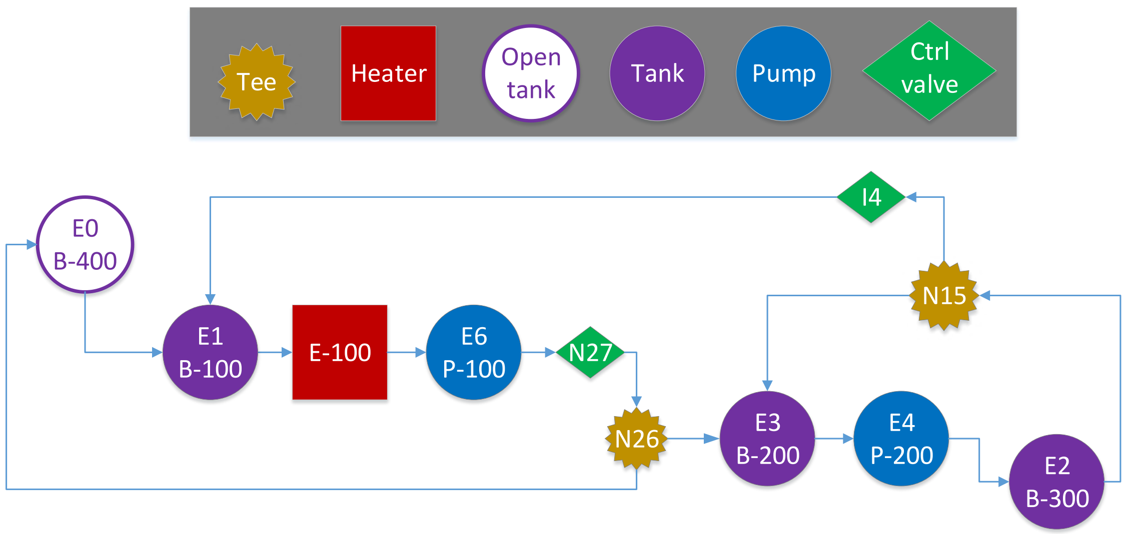


Figure 12. The result of processing the graph in Figure 11 with the algorithm in Figure 4. The tank with an internal heater is presented by adding the heater to the outgoing flow of the tank.

Table 2. Components (i.e., symbol in Balas®) of the steady state model

NodeName	Symbol
B-400	MDTank#1
Source1	Terminal in
B-100	MDTank#1
Source2	Terminal in
B-300	MDTank#1
Source3	Terminal in
B-200	MDTank#1
Source4	Terminal in
P-200	Pump
P-100	Pump
I4	Valve
ES-E100	Heater/cooler#1
N15	Splitter
N26	Splitter
N27	Valve

Table 3. Flows of the steady state model. The strings in 'Source' and 'Target' columns refer to symbol names in the 'NodeName' column of Table 2.

Source	SourcePort	Target	TargetPort
Source1	1	B-400	2
B-400	1	drain	0
Source2	1	B-100	2
B-100	1	drain	0
Source3	1	B-300	2
B-300	1	drain	0
Source4	1	B-200	2
B-200	1	drain	0
B-200	3	P-200	1
B-400	3	B-100	3
I4	2	B-100	4
N15	1	I4	1
B-300	3	N15	3
N15	2	B-200	4
P-100	2	N27	1
N27	2	N26	3
N26	1	B-200	5
N26	2	B-400	4
ES-E100	2	P-100	1
P-200	2	B-300	4
B-100	5	ES-E100	1

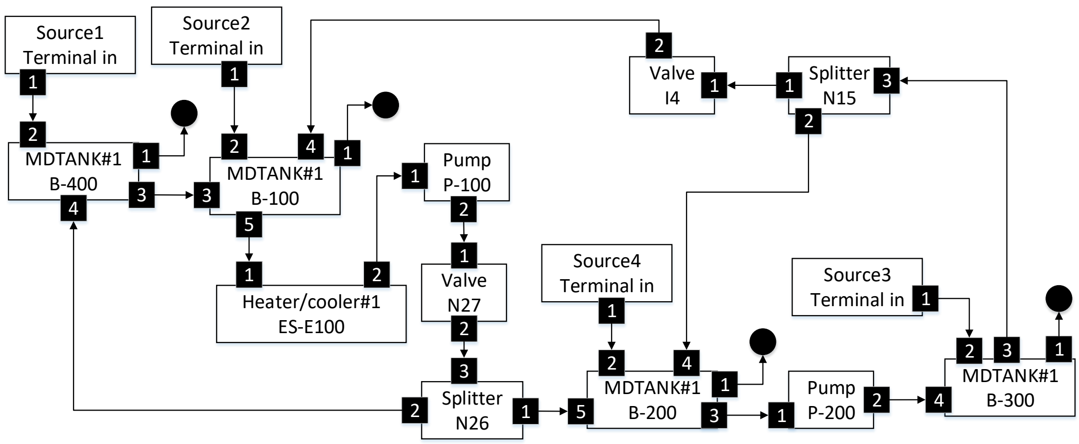


Figure 13. Visualization of the Balas® model specified in Tables 2 and 3.

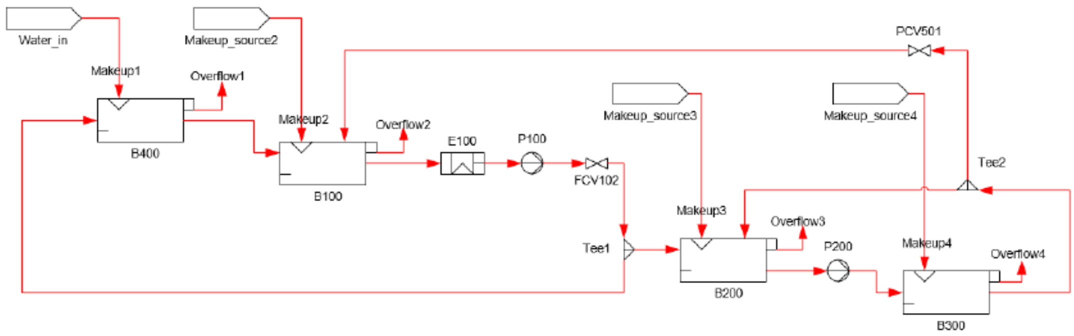


Figure 14. Imported flowsheet of the steady state model.

The simplified calculation module can be used to simulate a storage tank with several inflows and one outflow. The module mixes the inflows together and provides one outflow. This calculation module is used for simulating tanks that during steady state do not have any makeup flow or overflow but rather only a flow-through. In the case process, during steady state, the tanks "B200" and "B300" are such flow-through tanks and can for simplicity be simulated using the more simplified calculation module. These changes of the tank calculation modules (as well as the visual symbol of the tank) for tanks "B200" and "B300" were done manually based on the expertise of the modeler. Figure 15 shows the result of the manual finalizations of the flowsheet presented in Figure 14.

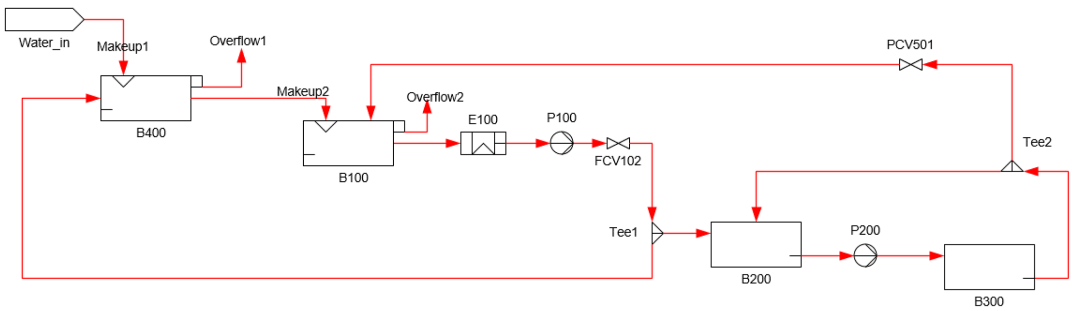


Figure 15. Final flowsheet of the steady state model resulting from expert manual finalizations of the model in Figure 14.

It is very common that the modeler changes the calculation modules for tanks (and the visual symbols presenting the tanks) during the iterative simulation work. In real processes, e.g., in paper machines, the water circuits are very complex connecting several tanks together. The makeup and overflow streams of the tanks are connected across. Some tanks may have a makeup stream from the freshwater system and some overflow streams directed to the sewer. To make a rule for such water circuits would require studying the entire circuit. This is time consuming. Instead, having a rule to model each tank as a complicated one and then later manually simplify the system is faster.

Step 6 of the methodology in Section 3.1 involves the initialization of the model by selecting the chemical components and the calculation modules for the symbols modelling the process equipment in Figure 15.

The initialization of the model starts with selecting manually the chemical components present in the process. In the simulation model, the chemical compositions and the conditions (T,p) of each feed stream must be defined. The chemical compositions of other streams (internal and products) are calculated automatically when the model is run. In the case process, water was the only component present in the feed stream ("Water in").

After defining the chemical components, the initialization of the model continues with selecting the calculation modules. Each symbol may have one to more calculation modules, which are selected manually in the simulation tool from a drop-down list. For example, the symbol simulating a heater may have a calculation module for either defining the outlet temperature or the thermal duty of the heater. The symbol simulating a valve may have a module for either defining the outlet pressure or the flow through the valve. The splitters may have a module for either defining the true mass flow (kg/s) of the first outlet or the share of the flow to the first outlet. In the case process, the modeler selected the calculation modules manually based on his or her expertise. For example, the calculation module for the heater "E100" was selected to be "the defined outlet temperature", since the temperature of the water circulating in the case process was known. The calculation module for the valve "FCV102" was selected to be "fixed flow". This setpoint value sets the amount of the circulating water in the system.

It is typical that during the iterative simulation work, the modeler changes the calculation modules of the symbols depending on what kind of input data (e.g., temperature, thermal duty, flow, etc.) is available.

Step 7 of the methodology in Section 3.1 involves parameterization of the calculation modules for the symbols in Figure 15. The selected calculation module determines the set of input values that are needed to parametrize the module. Possible parameters are, for example, exit pressure (kPa), exit temperature (°C), flow (kg/s), pump efficiency (%) or share of flow to specified stream in a junction. At this point of the research work, the parametrization of the modules was done manually based on the input data available for the case process.

After parametrization, the model can be finally run. Figure 16 shows the results of the simulation model describing the case process. At steady state, there is water circulating through the main line of the system, namely through tanks "B100", "B200", and "B300". Since no leaks are assumed, both the makeup streams have zero flow.

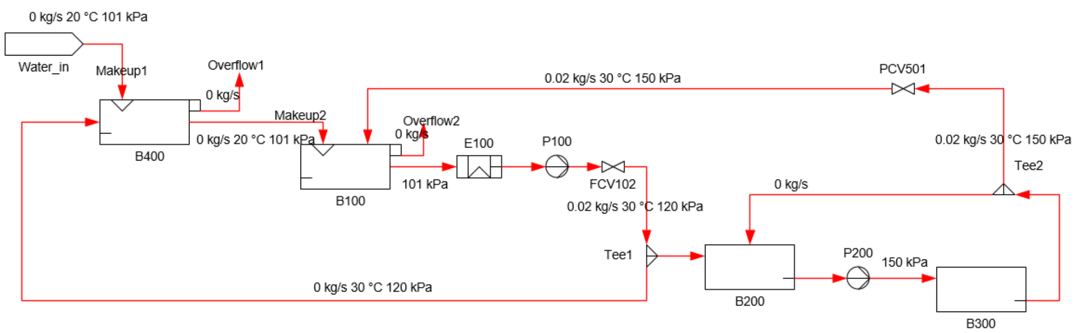


Figure 16. Simulations results of the case process.

5. Results

The case study has served as a proof-of-concept (POC) to validate the proposed algorithms. The great majority of manual engineering work was automated with respect to generating the flowsheet of the steady state model. It was discovered that all of the work that could not be automated involved the application of expert reasoning that could not readily be captured as general-purpose rules or algorithms. Thus, it was found that the developed approach is not expected to replace the human expert, but rather has potential to increase the engineering productivity of the expert. The findings are insufficient for the purpose of drawing any conclusions about the correctness or extensibility of the proposed algorithms for industrial grade processes. However, the findings about the extent of engineering work that could be automated for this case study indicate that the algorithms are ready for further research in the context of significantly more complex processes.

6. Discussion

The target of the paper was to achieve a POC for the automatic generation of a steady state model. The case process selected for the POC is simple and contains only simple unit operations such as tanks, pumps, and valves that can be modelled in the selected simulation tool with one single symbol. As a result, the rules for a one-to-one mapping from an intermediate graph to a steady state model presented in Table 1 are very simple. For chemical processes with more complicated unit operations, such as distillation columns, evaporators or extractors, the rules are longer since a distillation column or a liquid–liquid extractor, for example, are modelled by combining several symbols in series or parallel instead of having only one symbol. Even though the rule is longer, the same approach is applicable as with the simpler rules presented in this paper. However, further research is needed to define a set of rules to cover the most common chemical unit operations.

If the equipment is not available in the library of the simulation software, it can be modelled by creating a custom combination of equipment available in the library. In the case process, the rules are one-to-one mapping (from one piece of equipment to one symbol in the steady state model). In the said further research on chemical unit operations, several-to-one and/or one-to-several mapping rules are expected. Commercial simulation software may have an emphasis on a certain chemical process technology. For example, the simulation software AspenPlus is powerful for modelling unit operations based on phase separation whereas Balas is designed for modelling and simulating of paper processes. Thus, in AspenPlus there is no single symbol for a headbox of a paper machine and vice versa, in Balas, there is no single symbol for a distillation column. The rule for describing any unit operation is always simulation software specific. In the absence of standardization in the area of steady state simulation tools, one direction for further research would be the development of rules for simulator-to-simulator transfer.

The methodology presented in this paper assumes as its starting point that the relevant information has been extracted from a P&ID into a graph format. For many brownfield plants, the P&ID is a raster graphics image obtained by scanning a paper diagram. In newer plants a digital P&ID from a CAD tool may be available. The latter scenario applies to our case study. Thus, the methodology is general for all kinds of plants. However, it has been especially designed to work on the limited information at a brownfield plant. In the case of raster graphics P&IDs, the quality of the results obtained by this methodology depend additionally on the quality of the P&ID information extraction solution. The availability of recent publications in this area by several research groups, referenced in bullet 1 of the numbered list in Section 3.1, is an indication that efforts are underway to further advance the quality of information extraction from brownfield design documents.

7. Conclusion and Further Work

7.1. Limitations

Flowsheet generation is always a necessary step when building a model. The complexity of generating a flowsheet does not depend on the complexity of the configuration of the process. The flowsheet does not describe the chemical and physical phenomena occurring in the process (i.e., model components and their reactions). The flowsheet describes the connection between the process equipment (i.e., process configuration). The solution presented in this paper is focused on the flowsheet generation. This article has not targeted the information needed to describe thermo-hydraulic or chemical phenomena. To overcome these limitations, it is necessary to (i) select suitable calculation modules, (ii) parametrize the model based on available data and information, and (iii) set the parameter values to the model. The parameters include (i) unit operation input parameters, (ii) feed stream composition and state, (iii) design specification, (iv) solver parameters, and (v) thermodynamic model parametrization. With this information, the flowsheet is supplemented with adequate information of the phenomena and it is possible to simulate the process with the model.

7.2. Summary of Results and Further Work

In this paper, a 7-step methodology was proposed for the generation of steady state digital twins for process plants. Related works were positioned along the steps 1–2, so the focus of the paper was on steps 3–7. The findings and topics of further work for these steps are discussed next.

In our case study, the result of steps 3 and 4 was an automatically generated flowsheet of a steady state model that required only minor manual changes by an expert modeler. Specifically, the following changes were made to two of the tanks: changing the type of the tank to another type of tank from the library of the steady state modelling tool and reconnecting the makeup flows of these tanks. It may be concluded that the generation significantly reduced the manual modelling effort and that the methodology is ready for further research on larger and more complicated processes.

The rationale for the manual changes done in step 5 was presented in detail. The modelling decisions related to a makeup flow of a tank required the consideration of several parts of the process upstream of the said tank. The general formulation of such modelling decisions as rules is a nontrivial problem. However, one direction of further research would be the formulation and implementation of such rules, and validating them across a wide range of case studies.

In step 6, the modelling decisions related to selecting calculation modules were discussed. It was noted that the decisions depend on the properties of incoming flows, which in turn depend on how other parts of the process were modelled. The automation of this work was left for further research.

In step 7, concrete examples of steady state model parameterization were given, and the case model was parameterized manually based on known typical operating parameters of the process. In further work, the developed toolchain could be integrated to the process automation system and its history database in order to retrieve recent sensor values and to use them to automatically parameterize the steady state model. It is proposed that such a capability for automatic parameterization would turn a steady state simulation model to a steady state digital twin. It is notable that there is a lack of research specifically about digital twins based on steady state models, so there is no established definition for a steady state digital twin. Significant further research questions arise related to the development of the automatic parameterization capability, so it is not only an industrial information integration task. Knowledge about the recent operating conditions of the process is required to select and preprocess a suitable time period of recent process history, in order to parameterize a steady state model that will be relevant for answering the specific questions related to the unique retrofit project at hand.

Author Contributions: Conceptualization, S.S., L.S., and M.A.; Data curation, S.S. and L.S.; Formal analysis, S.S. and L.S.; Funding acquisition, S.S. and E.H.; Investigation, S.S.; Methodology, S.S. and L.S.; Project administration, S.S. and E.H.; Resources, S.S., L.S., and V.V.; Software, S.S. and A.V.; Supervision, S.S., E.H., and V.V.; Validation, S.S., L.S., and A.V.; Visualization, S.S. and M.A. Writing—original draft, S.S., L.S., M.A., and A.V.; Writing—review & editing, S.S., L.S., M.A., E.H., and V.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Business Finland grants 3915/31/2019 and 4153/31/2019.

Conflicts of Interest: The authors declare no conflict of interest. The funder had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Pinto-Varela, T.; Barbosa-Póvoa, A.P.F.D.; Carvalho, A. Sustainable batch process retrofit design under uncertainty—An integrated methodology. *Comput. Chem. Eng.* **2017**, *102*, 226–237. [CrossRef]
2. Wang, B.; Klemeš, J.J.; Varbanov, P.S.; Chin, H.H.; Wang, Q.-W.; Zeng, M. Heat exchanger network retrofit by a shifted retrofit thermodynamic grid diagram-based model and a two-stage approach. *Energy* **2020**, *198*, 117338. [CrossRef]
3. Min, K.-J.; Binns, M.; Oh, S.-Y.; Cha, H.-Y.; Kim, J.-K.; Yeo, Y.-K. Screening of site-wide retrofit options for the minimization of CO_2 emissions in process industries. *Appl. Therm. Eng.* **2015**, *90*, 335–344. [CrossRef]
4. Faria, D.C.; Bagajewicz, M.J. Profit-based grassroots design and retrofit of water networks in process plants. *Comput. Chem. Eng.* **2009**, *33*, 436–453. [CrossRef]
5. Wen, M.; Wu, Q.; Li, G.; Wang, S.; Li, Z.; Tang, Y.; Xu, L.; Liu, T. Impact of ultra-low emission technology retrofit on the mercury emissions and cross-media transfer in coal-fired power plants. *J. Hazard. Mater.* **2020**, *396*, 122729. [CrossRef] [PubMed]

6. Sierla, S.; Azangoo, M.; Vyatkin, V.; Fay, A.; Papakonstantinou, N. Integrating 2D and 3D Digital Plant Information towards Automatic Generation of Digital Twins. In Proceedings of the 29th IEEE International Symposium on Industrial Electronics, Delft, The Netherlands, 17–19 June 2020.
7. Chen, B.; Wan, J.; Shu, L.; Li, P.; Mukherjee, M.; Yin, B. Smart Factory of Industry 4.0: Key Technologies, Application Case, and Challenges. *IEEE Access* **2018**, *6*, 6505–6519. [CrossRef]
8. Schluse, M.; Priggemeyer, M.; Atorf, L.; Roßmann, J.; Romann, J. Experimentable Digital Twins—Streamlining Simulation-Based Systems Engineering for Industry 4.0. *IEEE Trans. Ind. Inform.* **2018**, *14*, 1722–1731. [CrossRef]
9. Martinez, G.S.; Karhela, T.A.; Ruusu, R.J.; Sierla, S.; Vyatkin, V. An Integrated Implementation Methodology of a Lifecycle-Wide Tracking Simulation Architecture. *IEEE Access* **2018**, *6*, 15391–15407. [CrossRef]
10. Martinez, G.S.; Sierla, S.; Karhela, T.; Lappalainen, J.; Vyatkin, V. Automatic Generation of a High-Fidelity Dynamic Thermal-Hydraulic Process Simulation Model from a 3D Plant Model. *IEEE Access* **2018**, *6*, 45217–45232. [CrossRef]
11. Martínez, G.S.; Sierla, S.; Karhela, T.; Vyatkin, V. Automatic Generation of a Simulation-Based Digital Twin of an Industrial Process Plant. In Proceedings of the 44th Annual Conference of the IEEE Industrial Electronics Society IECON 2018, Washington, DC, USA, 21–23 October 2018; pp. 3084–3089. [CrossRef]
12. Arroyo, E.; Hoernicke, M.; Rodríguez, P.; Fay, A. Automatic derivation of qualitative plant simulation models from legacy piping and instrumentation diagrams. *Comput. Chem. Eng.* **2016**, *92*, 112–132. [CrossRef]
13. Shellshear, E.; Berlin, R.; Carlson, J.S. Maximizing Smart Factory Systems by Incrementally Updating Point Clouds. *IEEE Comput. Graph. Appl.* **2015**, *35*, 62–69. [CrossRef] [PubMed]
14. Emerson. Understanding and Applying Simulation Fidelity to the Digital Twin. White Paper 2018. Available online: emerson.com/documents/automation/understanding-applying-simulation-fidelity-to-digital-twin-en-5079366 (accessed on 2 October 2020).
15. Matzopoulos, M. Dynamic Process Modeling: Combining Models and Experimental Data to Solve Industrial Problems. In *Process Systems Engineering: Volume 7 Dynamic Process Modeling*; Georgiadis, M.C., Banga, J.R., Pistikopoulos, E.N., Eds.; Wiley-VCH: Weinheim, Germany, 2010; pp. 1–33. [CrossRef]
16. Leiviskä, K. *Simulation in Pulp and Paper Industry*; University of Oulu, Control Engineering Laboratory: Oulu, Finland, 1996; Report A, 2; pp. 1–58.
17. Blanco, A.; Dahlquist, E.; Kappen, J.; Manninen, J.; Negro, C.; Ritala, R. Use of modelling and simulation in the pulp and paper industry. *Math. Comput. Model. Dyn. Syst.* **2009**, *15*, 409–423. [CrossRef]
18. Bezzo, F.; Bernardi, R.; Cremonese, G.; Finco, M.; Barolo, M. Using Process Simulators for Steady-State and Dynamic Plant Analysis. *Chem. Eng. Res. Des.* **2004**, *82*, 499–512. [CrossRef]
19. Enaasen, N.; Tobiesen, A.; Kvamsdal, H.M.; Hillestad, M. Dynamic Modeling of the Solvent Regeneration Part of a CO_2 Capture Plant. *Energy Procedia* **2013**, *37*, 2058–2065. [CrossRef]
20. Măluțan, T.; Măluţan, C. Simulation of Processes in Papermaking by WinGEMS Software. *Environ. Eng. Manag. J.* **2013**, *12*, 1645–1647. [CrossRef]
21. Turon, X.; Labidi, J.; Paris, J. Simulation and optimisation of a high grade coated paper mill. *J. Clean. Prod.* **2005**, *13*, 1424–1433. [CrossRef]
22. Cardoso, M.; De Oliveira, K.D.; Costa, G.A.A.; Passos, M.L. Chemical process simulation for minimizing energy consumption in pulp mills. *Appl. Energy* **2009**, *86*, 45–51. [CrossRef]
23. Atkins, M.; Morrison, A.; Walmsley, M.; Riley, J. WinGEMS Modelling and Pinch Analysis of a Paper Machine for Utility Reduction. *Appita J.* **2010**, *63*, 281–287.
24. Jönsson, J.; Ruohonen, P.; Michel, G.; Berntsson, T. The potential for steam savings and implementation of different biorefinery concepts in Scandinavian integrated TMP and paper mills. *Appl. Therm. Eng.* **2011**, *31*, 2107–2114. [CrossRef]
25. Clement, S.; Gouiller, A.; Ottenio, P.; Nivelon, S.; Huber, P.; Nortier, P. Speciation and supersaturation model in papermaking streams. *Process. Saf. Environ. Prot.* **2011**, *89*, 67–73. [CrossRef]
26. Huber, P.; Nivelon, S.; Ottenio, P.; Nortier, P. Coupling a Chemical Reaction Engine with a Mass Flow Balance Process Simulation for Scaling Management in Papermaking Process Waters. *Ind. Eng. Chem. Res.* **2012**, *52*, 421–429. [CrossRef]
27. Kangas, P.; Kaijaluoto, S.; Määttänen, M. Evaluation of future pulp mill concepts—Reference model of a modern Nordic kraft pulp mill. *Nord. Pulp Pap. Res. J.* **2014**, *29*, 620–634. [CrossRef]

28. Barbera, E.; Menegon, S.; Banzato, D.; D'Alpaos, C.; Bertucco, A. From biogas to biomethane: A process simulation-based techno-economic comparison of different upgrading technologies in the Italian context. *Renew. Energy* **2019**, *135*, 663–673. [CrossRef]
29. Kautto, J.; Realff, M.J.; Ragauskas, A.J. Design and simulation of an organosolv process for bioethanol production. *Biomass Convers. Biorefinery* **2013**, *3*, 199–212. [CrossRef]
30. Søtoft, L.F.; Rong, B.-G.; Christensen, K.; Norddahl, B. Process simulation and economical evaluation of enzymatic biodiesel production plant. *Bioresour. Technol.* **2010**, *101*, 5266–5274. [CrossRef]
31. Cheah, K.W.; Yusup, S.; Singh, H.K.G.; Uemura, Y.; Lam, H.L.; Wai, C.K. Process simulation and techno economic analysis of renewable diesel production via catalytic decarboxylation of rubber seed oil—A case study in Malaysia. *J. Environ. Manag.* **2017**, *203*, 950–961. [CrossRef]
32. Barbosa, L.D.S.N.S.; Hytönen, E.; Vainikka, P. Carbon mass balance in sugarcane biorefineries in Brazil for evaluating carbon capture and utilization opportunities. *Biomass Bioenergy* **2017**, *105*, 351–363. [CrossRef]
33. Hytönen, E.; Stuart, P.R. Biofuel Production in an Integrated Forest Biorefinery—Technology Identification under Uncertainty. *J. Biobased Mater. Bioenergy* **2010**, *4*, 58–67. [CrossRef]
34. Nabgan, B.; Abdullah, T.A.T.; Nabgan, W.; Ahmad, A.; Saeh, I.; Moghadamian, K. Process Simulation for Removing Impurities from Wastewater Using Sour Water 2-Strippers system via Aspen Hysys. *Chem. Prod. Process. Model.* **2016**, *11*, 315–321. [CrossRef]
35. Miltner, A.; Wukovits, W.; Pröll, T.; Friedl, A. Renewable hydrogen production: A technical evaluation based on process simulation. *J. Clean. Prod.* **2010**, *18*, S51–S62. [CrossRef]
36. Zhang, Y.; Cruz, J.; Zhang, S.; Lou, H.H.; Benson, T.J. Process simulation and optimization of methanol production coupled to tri-reforming process. *Int. J. Hydrogen Energy* **2013**, *38*, 13617–13630. [CrossRef]
37. Michaux, B.; Rudolph, M.; Reuter, M.A.; Reuter, M.A. Study of process water recirculation in a flotation plant by means of process simulation. *Miner. Eng.* **2020**, *148*, 106181. [CrossRef]
38. McNulty, M.J.; Gleba, Y.; Tusé, D.; Hahn-Löbmann, S.; Giritch, A.; Nandi, S.; McDonald, K.A. Techno-economic analysis of a plant-based platform for manufacturing antimicrobial proteins for food safety. *Biotechnol. Prog.* **2019**, *36*, e2896. [CrossRef] [PubMed]
39. Bon, J.; Clemente, G.; Váquiro, H.; Mulet, A. Simulation and optimization of milk pasteurization processes using a general process simulator (ProSimPlus). *Comput. Chem. Eng.* **2010**, *34*, 414–420. [CrossRef]
40. Koulamas, C.; Kalogeras, A.P. Cyber-Physical Systems and Digital Twins in the Industrial Internet of Things. *Computer* **2018**, *51*, 95–98. [CrossRef]
41. Tao, F.; Zhang, M. Digital Twin Shop-Floor: A New Shop-Floor Paradigm towards Smart Manufacturing. *IEEE Access* **2017**, *5*, 20418–20427. [CrossRef]
42. Zhang, H.; Liu, Q.; Chen, X.; Zhang, D.; Leng, J. A Digital Twin-Based Approach for Designing and Multi-Objective Optimization of Hollow Glass Production Line. *IEEE Access* **2017**, *5*, 26901–26911. [CrossRef]
43. Wan, J.; Tang, S.; Li, D.; Imran, M.; Zhang, C.; Liu, C.; Pang, Z. Reconfigurable Smart Factory for Drug Packing in Healthcare Industry 4.0. *IEEE Trans. Ind. Inform.* **2018**, *15*, 507–516. [CrossRef]
44. Schmidt, N.; Lüder, A. The Flow and Reuse of Data: Capabilities of Automation ML in the Production System Life Cycle. *IEEE Ind. Electron. Mag.* **2018**, *12*, 59–63. [CrossRef]
45. Hartmann, B.; Török, S.; Börcsök, E.; Groma, V.O. Multi-objective method for energy purpose redevelopment of brownfield sites. *J. Clean. Prod.* **2014**, *82*, 202–212. [CrossRef]
46. Sørensen, D.; Brunoe, T.D.; Nielsen, K. Brownfield Development of Platforms for Changeable Manufacturing. *Procedia CIRP* **2019**, *81*, 986–991. [CrossRef]
47. Illa, P.K.; Padhi, N. Practical Guide to Smart Factory Transition Using IoT, Big Data and Edge Analytics. *IEEE Access* **2018**, *6*, 55162–55170. [CrossRef]
48. Barth, M.; Fay, A. Automated generation of simulation models for control code tests. *Control. Eng. Pract.* **2013**, *21*, 218–230. [CrossRef]
49. Stojanovic, N.; Milenovic, D. Data-driven Digital Twin approach for process optimization: An industry use case. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 4202–4211.
50. Makarov, V.; Frolov, Y.; Parshina, I.S.; Ushakova, M. The Design Concept of Digital Twin. In Proceedings of the 2019 Twelfth International Conference "Management of large-scale system development" (MLSD), Moscow, Russia, 1–3 October 2019; pp. 1–4.

51. Kychkin, A.; Nikolaev, A. IoT-based Mine Ventilation Control System Architecture with Digital Twin. In Proceedings of the 2020 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), Sochi, Russia, 18–22 May 2020; pp. 1–5.
52. Barth, M.; Strube, M.; Fay, A.; Weber, P.; Greifeneder, J. Object-oriented engineering data exchange as a base for automatic generation of simulation models. In Proceedings of the 2009 35th Annual Conference of IEEE Industrial Electronics, Porto, Portugal, 3–5 November 2009; pp. 2465–2470.
53. Campos, J.G.; López, J.S.; Quiroga, J.I.A.; Seoane, A.M.E. Automatic generation of digital twin industrial system from a high level specification. *Procedia Manuf.* **2019**, *38*, 1095–1102. [CrossRef]
54. Sierla, S.; Azangoo, M.; Vyatkin, V. Generating an Industrial Process Graph from 3D Pipe Routing Information. In Proceedings of the 25th IEEE International Conference on Emerging Technologies and Factory Automation, ETFA 2020, Vienna, Austria, 8–11 September 2020.
55. Wen, R.; Tang, W.; Su, Z. Topology based 2D engineering drawing and 3D model matching for process plant. *Graph. Model.* **2017**, *92*, 1–15. [CrossRef]
56. Rantala, M.; Niemistö, H.; Karhela, T.; Sierla, S.; Vyatkin, V. Applying graph matching techniques to enhance reuse of plant design information. *Comput. Ind.* **2019**, *107*, 81–98. [CrossRef]
57. Son, H.; Kim, C.; Kim, C. 3D reconstruction of as-built industrial instrumentation models from laser-scan data and a 3D CAD database based on prior knowledge. *Autom. Constr.* **2015**, *49*, 193–200. [CrossRef]
58. Lee, J.; Son, H.; Kim, C.; Kim, C. Skeleton-based 3D reconstruction of as-built pipelines from laser-scan data. *Autom. Constr.* **2013**, *35*, 199–207. [CrossRef]
59. Papakonstantinou, N.; Karttunen, J.; Sierla, S.; Vyatkin, V. Design to automation continuum for industrial processes: ISO 15926–IEC 61131 versus an industrial case. In Proceedings of the 2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Zaragoza, Spain, 10–13 September 2019; pp. 1207–1212.
60. Sinha, A.; Bayer, J.; Bukhari, S.S. Table Localization and Field Value Extraction in Piping and Instrumentation Diagram Images. In Proceedings of the 2019 International Conference on Document Analysis and Recognition Workshops (ICDARW), Sydney, Australia, 20–25 September 2019; Volume 1, pp. 26–31.
61. Nurminen, J.K.; Rainio, K.; Numminen, J.P.; Syrjänen, T.; Paganus, N.; Honkoila, K. Object Detection in Design Diagrams with Machine Learning. In *Advances in Intelligent Systems and Computing*; Springer: Cham, Germany, 2020; Volume 977, pp. 27–36. [CrossRef]

applied sciences

MDPI

Article

Digital Twin for Monitoring Ergonomics during Manufacturing Production

Alessandro Greco, Mario Caterino *, Marcello Fera and Salvatore Gerbino

Department of Engineering, University of Campania Luigi Vanvitelli, via Roma 29, 81031 Aversa, Italy; alessandro.greco@unicampania.it (A.G.); marcello.fera@unicampania.it (M.F.); salvatore.gerbino@unicampania.it (S.G.)

* Correspondence: mario.caterino@unicampania.it; Tel.: +39-081-50-10-318

Received: 30 September 2020; Accepted: 30 October 2020; Published: 2 November 2020

check for updates

Abstract: Within the era of smart factories, concerning the ergonomics related to production processes, the Digital Twin (DT) is the key to set up novel models for monitoring the performance of manual work activities, which are able to provide results in near real time and to support the decision-making process for improving the working conditions. This paper aims to propose a methodological framework that, by implementing a human DT, and supports the monitoring and the decision making regarding the ergonomics performances of manual production lines. A case study, carried out in a laboratory, is presented for demonstrating the applicability and the effectiveness of the proposed framework. The results show how it is possible to identify the operational issues of a manual workstation and how it is possible to propose and test improving solutions.

Keywords: ergonomics; manufacturing; production process; Digital Twin

1. Introduction

Ergonomic issues represent one of the main factors characterizing the manufacturing working environment. Indeed, in addition to designing a production system according to an ergonomic approach, the working activities need to be continuously monitored, especially when the volume of production varies, causing changes in cycle time and, consequently, changes in working tasks and workload.

Going more into detail, biomechanical overload represents one of the main risk factors in a manufacturing environment and is a possible source of Musculo-Skeletal Disorders (MSDs). MSDs consist in lesions or alterations of muscles, nerves, tendons and joints and, as demonstrated by numerous studies in literature and international standards, they are caused by prolonged exposure to awkward working postures, exerted forces, Material Manual Handling (MMH) and repetitive actions [1,2].

Over the years, the high incidence of the MSDs related to biomechanical overload led to the development of numerous risk assessment methods, principally observation-based, as collected by Takala et al. [3]. The application of risk assessment methods is mandatory during normal production to monitor working activities whenever changes are made to production volume or to working tasks. However, since the methods are observational, mostly based on the compilation of specific check-lists, the evaluation procedure is time-consuming and, additionally, strictly subjective, so it does not guarantee the repeatability of measurements.

For these reasons, in the era of Industry 4.0, the need to develop numerical methodologies and measurement devices has become a priority for companies wishing to have efficient and safe production systems. In particular, among the enabling technologies, the integrated use of Industrial Internet of Things (IIoT) and simulation tools allows for monitoring working performance, enhancing the control of the process and a near-real time management of the whole production line, in terms of balancing and ergonomics [4,5]. However, the literature dealing with the use of Industry 4.0 concept and tools for Ergonomics is poor, as also pointed out by Kadir et al. [6] in a deep review. Most researches are documented in conference proceedings, although the topic is very actual with wide application possibilities.

This research was born with the aim of creating tools that can support experts in ergonomic screening, ensuring accuracy and repeatability of measurements. From this point of view, the DT, based on the integration of IIoT and simulation, can represent the best solution as it allows exploiting the computational capability by using numerical models in which real data are implemented.

IIoT includes objects or devices (sensors, actuators, mobiles, etc.) that are able to interact and communicate between them, via internet protocols, and, thanks to the implementation of specific algorithms, they are able to carry out measurements and make decisions autonomously in order to manage machines and production systems [7]. Among these technologies, for ergonomics purposes, wearable devices are crucial since they can be provided with sensors capable to measure parameters related to humans (postures, forces, muscular activity, etc.). Widespread in recent years, wearable motion tracking devices have seen a massive introduction in the factory environment. Thanks to their good accuracy and low invasiveness, they allow acquiring motion data during normal working activities and analyzing them in order to evaluate ergonomics indexes related to the biomechanical load due to working postures [8–10].

In addition to IIoT, a manufacturing scenario can be fully reproduced in a 3D environment, giving the possibility to simulate the operations of a real working process [11]. This makes feasible, since the design phase, choosing the correct solutions, knowing in advance the performance of the line. The implementation of real data, collected by IIoT devices, in a simulation scenario realizes the so-called Digital Twin (DT).

Manual operations are still dominant in complex manufacturing systems [12], hence the working tasks are affected by high variability. By means of DT, it is interesting to investigate the possibility to implement real data into a simulation in order to assess ergonomics indexes.

In literature, several researches deal with digital human modeling and simulation for assessing ergonomics indexes. Caputo et al. [13,14] defined experimental/numerical procedures to evaluate the ergonomics of working activities by using, separately, inertial motion tracking system and simulation. Experimental data were used in Caputo et al. [13], for validating the numerical model, by comparing experimental and numerical results. The authors proposed a framework for preventively evaluating the ergonomic indexes (in particular, the Ergonomic Assessment Work-Sheet—EAWS) during the design phase of a new manufacturing line. In Caputo et al. [14], four ergonomic indexes, the same considered for the case study presented in Section 3, have been evaluated by means of numerical models for validating the design of a new workstation.

Makarova et al. [15] demonstrated that process parameters and ergonomics indexes can be investigated in a virtual environment. Case et al. [16] investigated the workers' ageing by implementing human capability data within a simulation environment. Tarallo et al. [17] proposed a computer-aided production control framework, which includes the use of digital human models, for implementing the principles of Industry 4.0 in manual working environments. The authors described how to monitor manual manufacturing processes by using a virtual simulation software (Siemens Tecnomatix Jack®) and an optical motion capture system (Microsoft Kinect®).

Sanjog et al. [18] used both physical and virtual ergonomics tool for assessing the ergonomics of an industrial shop-floor workstation. They pointed out that DHM and simulation could be very much beneficial for engineers/production supervisors/ergonomists to set the best design solution for a safe workstation.

However, these numerical models may not provide enough accurate results, since human motion is evaluated by inverse kinematic, making the movements sometimes unrealistic. So, implementing real data in Digital Human Models (DHM) should allow accurate simulation and assessment of production performance (ergonomics, working times, line efficiency, etc.).

Few significant studies have been documented in the technical literature about human simulation based on experimental data and most of them are relative to the control of human body for clinical purposes. Most of the research studies based on DT concern the improvement of manufacturing process and the product lifecycle management. In fact, there are no applications of human Digital Twin to assess the risk factors closely related to manual working activities.

Catarci et al. [19] provided a detailed literature about DTs and their modelling techniques, proposing a complex architecture for digital factories. Malik and Bilberg [20] presented a DT for investigating the performance of a human-robot collaboration work-cell, transferring the only data related to the robot, while Li et al. [21] developed an Augmented Reality (AR) application for the control of robots during a similar task, by implementing the DT of human hands by means of LeapMotion sensor and a Kinect V2 camera. Nikolakis et al. [22] used experimental data for the implementation of a DT for recognizing and simulating human activities.

Zheng et al. [23] studied the literature about DT technology for realizing a framework, aimed at product lifecycle management, based on three function modules: data storage, data processing and data mapping. Similarly, Ma et al. [24] proposed a framework based on DT to support the management of cyber-physical systems of production workshop, including product design and manufacturing. Aimed to improve the order management process, Kunath and Winkler [25] proposed a conceptual framework, based on DT, of a decision support system, which is able to find the best solution through simulating several scenarios. Instead, Havard et al. [26] combined DT and virtual reality in a co-simulation environment for assessing industrial workstations. They carried out a case study related to a human-robot collaborative workplace performing also ergonomic analysis.

This research aims to fill the gap in the literature about the use of human DT to evaluate ergonomics by proposing a methodology that supports ergonomists/occupational physicians/line managers in mapping the ergonomics risk for all the workstations of manufacturing environment.

There are two main motivations behind choosing a DT-based procedure for assessing the ergonomics. Firstly, it ensures that analyses are not affected by the ergonomist subjectivity, typical of the traditional (observational) techniques. Then, DT drastically reduces the computation times, thanks to algorithms able to quickly process data. So, since ergonomic analyses are significantly time consuming procedures, which could require hours for filling-in the spreadsheets, it is presumable saving a lot of time and then costs.

This paper moves from a previous research [4], in which the authors proposed a novel methodological framework, based on the implementation of a DT, for carrying out near real time analyses about the performance of manufacturing production lines, in terms of working times and balancing. The core of the research in [4] was related to human motion data collection and transferring for reproducing the real production in a simulation scenario. Such model is able to supply output data useful to the evaluation of the desired line performances, besides representing a predictive model for the behavior of the line itself following possible modifications.

Herein, the methodological framework has been modified and adapted in order to evaluate the worker's performance in terms of ergonomics, investigating the possible causes of risk of injury due to biomechanical overload: working postures, exerted forces, Material Manual Handling and repetitive actions. This procedure represents an innovation for the ergonomic screening of production lines that, currently, is still mainly performed by observational techniques and is a highly time consuming process. In fact, typically the ergonomist observes and records the work activity, estimates the index calculation parameters, fills in the checklists and evaluates the risk index. This process may require a significant amount of time, in the order of hours, to estimate the indexes of each workstation. In addition, the use of experimental data and DT allows an objective analysis as well as ensuring the repeatability of measurements. This approach contributes to the transformation towards the so-called smart factory, in agreement with the principles of Industry 4.0.

A case study, aimed to demonstrate the effectiveness of the framework, has been set and performed at the Laboratory of Machine Design of the University of Campania Luigi Vanvitelli. Data have been acquired by means of a wearable inertial motion tracking system [27] and the DT has been implemented in the Tecnomatix Process Simulate software by Siemens®.

The reminder of the paper is organized as follows. Section 2 describes the methodological framework, aimed to support the ergonomic assessment of the investigated working activity. Section 3 describes the case study that investigates an assembly activity reproduced in laboratory. Section 4 presents the results analysis and the discussions, while Section 5 concludes the paper.

2. Methodological Framework

Figure 1 describes the methodological framework developed to investigate about ergonomics of manual working tasks in a manufacturing scenario.

As already mentioned, the methodological framework moves from that one described in Fera et al. [4], which was related specifically on production line performance evaluation, whereas in this paper it has been modified and adapted for assessing the ergonomic indexes during the real production.

The framework consists of seven steps:

1. **Theoretical ergonomic balancing**: it is known and defined during the design and engineering phase of the production line;
2. **Production**: the workstation is selected for the investigation;
3. **Data collection**: experimental data about movements, forces, etc. are collected in order to allow the ergonomic evaluation;
4. **Simulation**: data are transferred to a DHM and the working activities are reproduced. Depending on the used devices, the simulation can be performed in real/near-real time or as post-process;
5. **Ergonomic assessment**: output data from simulation are used to evaluate the desired ergonomic indexes. In case the low risk condition is not satisfied, it is necessary to investigate and figure out the critical issues;
6. **Proposal and testing of improving solution**: critical issues need to be solved for satisfying the low case condition requirement. A time-based simulation is used for numerically testing and approving the proposed improving solution;
7. **Continue the production**: once the assessment is completed or the workplace is changed according to the proposed solution, the production can continue. A loop closes the framework, as the ergonomic mapping of the line has to be done every time there are changes in the production volume.

In the next paragraphs of this section, each phase will be explained and characterized. Data collection and simulation phases are the same of those described in [4], and briefly summarized in the following.

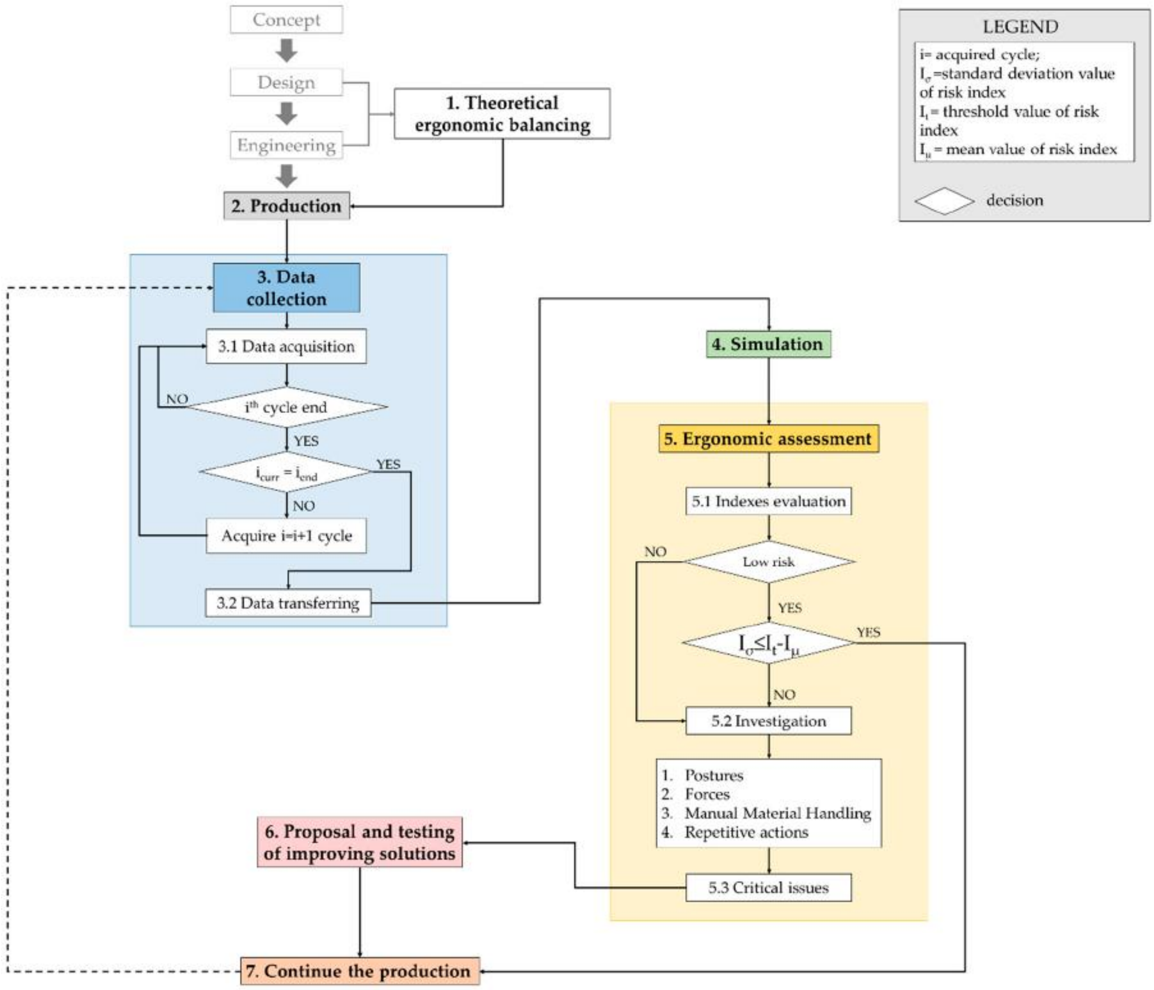


Figure 1. Methodological framework for monitoring ergonomics performance during the production process.

2.1. Data Collection

According to Fera et al. [4], data collection is a crucial step in order to apply the proposed approach. Specifically, data collection is related to the human body movements and it may be performed by means of both optical and non-optical motion capture based systems.

Together with the motion tracking system, it could be useful also to get additional data, such as the ones related to the exerted forces, by means of specific devices (force sensors, cyber gloves, etc.) to carry out the analysis.

It is important to define the number of cycles to be acquired during data acquisition session. This is required (see Section 3.1) in order to compute the basic statistical moments (mean and standard deviation) of ergonomic indexes, useful to assess whether a deep investigation is needed to evaluate critical ergonomic issues. The number of cycles to acquire during data collection phase strictly depends on the cycle time of the working task: in particular, the smaller the cycle time, the greater the number of acquisitions will be due to the increase of the variability of acquisitions [28].

Once the number of cycles to acquire have been defined, data acquisition session may start. If data acquisition is performed by means of wearable sensors, ethical requirement must be signed by the workers wearing them. Typically, a questionnaire about the usability of the system is also submitted to workers.

Acquired data are transferred to an appropriate software able to perform human simulations, such as Tecnomatix by Siemens® [29] or Delmia by Dassault Sistèmes® [30], and able to integrate such a data to the DHM, which accurately replicates the real workers' working tasks. Depending on the type of device used during the acquisition session, the way to transfer data to the software changes. In some cases, as shown later on in the case study, custom plugins and a re-sampling of data are necessary to integrate them in the simulation scenario.

2.2. Simulation

Step 4 in Figure 1 concerns simulation. Since it is focused on ergonomic assessment, simulation has to accurately replicate manual working tasks, according to the acquired data.

The steps [4] to follow in order to carry out the simulation are substantially four and involve: (i) virtual scenario setting, reproducing the same workplace layout; (ii) DHM creation, according to the anthropometric characteristics of the real worker; (iii) data implementation and operation refining when necessary (e.g., handling, picking or grasping an object, application of a force, etc.); (iv) run the simulation.

2.3. Ergonomics Assessment

Once the simulation is completed, numerical data are analyzed in order to perform the ergonomics assessment (step 5 in Figure 1). In a manufacturing scenario, as anticipated in Section 1, it is necessary to investigate the biomechanical overload, a source of injury risks, that is principally caused by working postures, exerted forces, MMH and repetitive actions with upper limbs. Many methods, tools and screening procedures have been developed over the years [3], some of which already implemented in different software codes for production process simulation. Many of them are according to the standards ISO 11,226 [31] and ISO 11228-1,2,3 [32–34] that regulate the whole procedure of occupational ergonomics monitoring.

After selecting the four risk indexes (one for each risk factor), numerical data are used for evaluating them (step 5.1 in Figure 1) for each acquired cycle. If the average values of the four indexes fall within the low risk area, the control phase is carried out by applying the following Equation (1) to each evaluated index:

$$I_{\sigma} \leq I_{t} - I_{\mu} \quad (1)$$

where I_{μ} and I_{σ} are the average and the standard deviation values of the evaluated index respectively, with respect to investigated working cycles, and I_t is the threshold value of the index for accessing the medium risk area.

If Equation (1) is satisfied, the framework suggests continuing the production (step 6 in Figure 1).

If none of the indexes fall within the risk area or if the Equation (1) is not satisfied, it is necessary furtherly investigating about the working task focusing on the critical factor.

This important step should be conducted by experienced ergonomists or by occupational physicians, since they have the appropriate know how for a proper identification and resolution of the critical issues.

Let us analyze in detail about the investigation (step 5.2 in Figure 1).

In case the working posture risk index does not fall within the low risk area, it is advisable to investigate, by observation or by studying the temporal history of postural angles, the postures assumed by the operator, identifying the sub-phases of the work cycle that mostly contribute to the value of the index.

About the exerted forces, if these exceed the maximum applicable value, as reported in specific tables by Snook and Ciriello [35], it is necessary to investigate the posture assumed during the force exertion, as well as the intensity and the duration of the application.

MMH is evaluated when the weight of the handled object is at least 3.5 kg; if the risk index exceeds the low risk area threshold, the investigations are different, depending on the kind of handling:

- for lifting operations, the attention has to be paid to the initial and final altimeters of the handling, as well as the type of gripping and frequency;

- for maintenance and carrying, the focus has to be on the distance travelled by the operator, together with the type of grip and frequency;
- for pushing/pulling actions, the investigation will be focused of the type and on the characteristics of the adopted cart, the distance covered by the worker and the frequency.

If the threshold value is exceeded by the index related to repetitive actions with upper limbs, the investigation will be focused on the number of technical actions, the possible awkward postures of the joints, the types of grip, the frequencies of actions and the recovery times.

Once the critical issues (step 5.3 in Figure 1) have been identified, it is advisable to discuss the possibility of making changes to the station layout or to the operations to be carried out.

After the discussion and testing of possible solutions, by means of numerical simulation, the ergonomist will deal with the decision-making process (step 6 in Figure 1) in order to improve the production process.

3. Case Study

Aiming to show the applicability and the effectiveness of the framework depicted in Figure 1, a case study is described here, related to a working task carried out in a laboratory, where a simple assembly task has been defined and performed. Figure 2 shows the working scenario reproduced in a simulation environment.

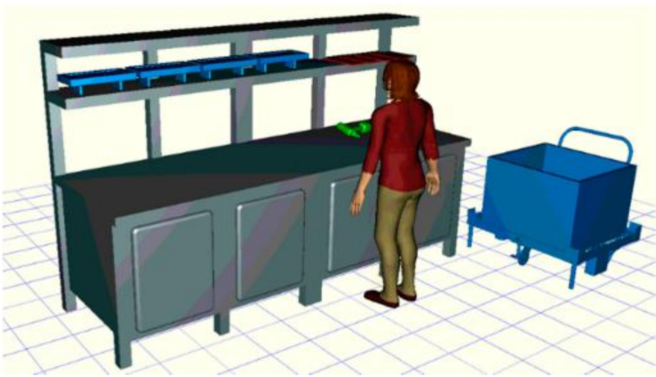


Figure 2. Workstation layout, reproduced in a simulation scenario.

As described in Figure 3, which shows both the real experiment and the Digital Twin, a female worker manually performs an assembly task of two components (labeled "1" and "2") made of steel. After picked up and positioned the components, these are joined by performing four screwings; then, the assembly is placed in a cart. A detailed description of each operation in reported in Table 1.

The assumed cycle time is 30 s, which includes 2 s of recovery. In addition, the work-shift duration has been assumed equal to 8 h, including 30 min breaks (10 min per break), equally distributed along the whole shift. A 60 min lunch-break has been scheduled at the middle of the shift.

Going into detail about the characteristics of the workstation, the components 1 and 2 have weight equal to 6 kg and 2 kg respectively, while the weight of the screwdriver is 2.5 kg.

About the screwings, the joints are made with M10 threaded bolts and the tightening torque is equal to 30 Nm.

The shelf where the two components are placed is 1400 mm high, while the height of the workbench is 900 mm. Finally, the assembly is positioned in the cart on a support plane 500 mm high.

The worker is P40 of the Italian female population [36], with a stature of 1550 mm and a weight of 45 kg.

In the following sections, the application of the methodological framework depicted in Figure 1 is applied.

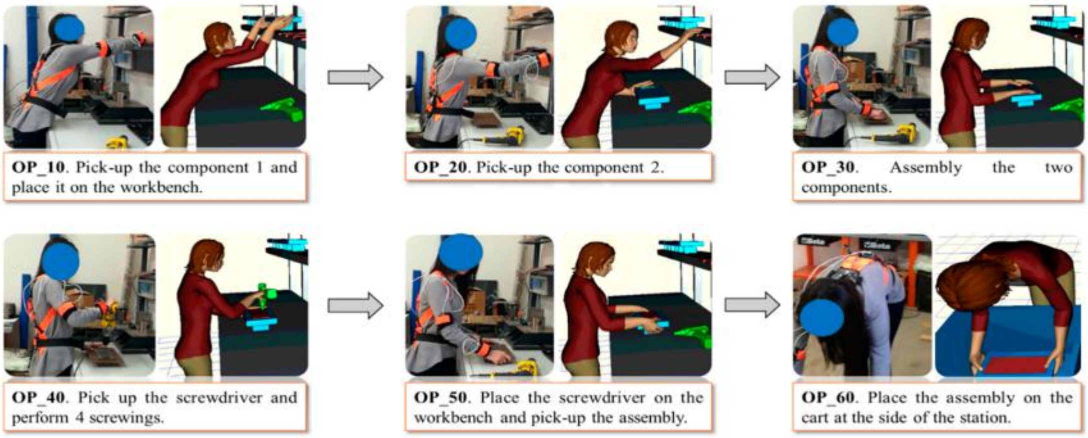


Figure 3. Tasks of the working activity. For each picture frame, on the left the real scenario and on the right the Digital Twin.

Table 1. Working cycle activities description.

Working Activities	
Operations	**Descriptions**
OP_10	The worker picks up the component 1, initially positioned on the shelf, with both hands, and then places it onto the workbench
OP_20	The worker picks up the component 2 from the shelf, with right hand, and places it above the component 1
OP_30	Matching the holes for the joining, the worker inserts and slightly turns four screws
OP_40	The worker picks up the screwdriver with right hand and performs four screwings
OP_50	The worker places the screwdriver on the workbench and picks up the assembly with both hands
OP_60	The worker reaches the cart at the side of the workstation and places the assembly

3.1. Data Collection

Motion data have been collected by using a wearable inertial motion tracking system developed at the Department of Engineering of the University of Campania Luigi Vanvitelli. The tracking system is composed by Inertial Movement Units and a sensor fusion algorithm, based on Extended Kalman Filter, has been developed for evaluating the attitude of body segments and, therefore, the posture angles related to the investigated working activity. Acquisition start and stop are managed by means of a mobile app used by an observer. A scheme of the motion tracking system is provided in Figure 4 and a full description of the algorithm can be found in [27]. This motion tracking system does not yet work in real time, so data transferring has been carried out after data processing. Indeed, the algorithm autonomously compiles CSV (Comma-Separated Values) files in which Euler angles, quaternions and posture angles are provided for each segment of the human body in each time frame.

The system is worn over the normal clothes by the worker (as in Figure 3) and, after a proper starting calibration of the sensors, the acquisition is run and the worker can normally perform the working tasks.

According to Giacomazzi [28], since the cycle time is about 30 s, motion data have been acquired for 60 consecutive working cycles (i = 60). Triggers have been manually introduced by the observer during the acquisition of data, in order to separate data related to consecutive working cycles. Figure 5 shows the trends of posture angles for trunk, elbow and arms along one working cycle.

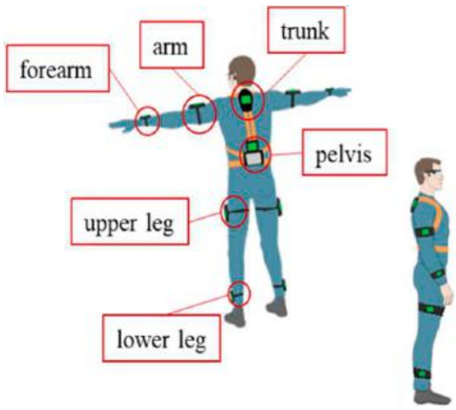


Figure 4. Wearable motion tracking system configuration.

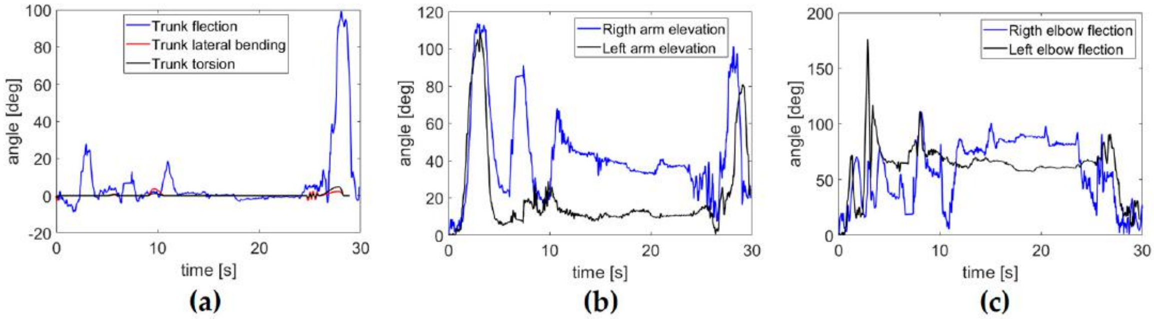


Figure 5. Posture angles trends over one working cycle for: trunk (**a**), arms (**b**) and elbows (**c**).

3.2. Data Transferring and Simulation

Data transferring is strongly influenced by the software chosen to perform simulation. Among commercial software for human simulation, most of them have interfaces that enable them to directly connect external supported devices, such as Kinect® or XSens®, to the simulation environment. Alternatively, several software packages allow implementing customized plugins to read data from external devices. To simulate the proposed working activity the software Tecnomatix Process Simulate by Siemens®, version 15.0.1, was used. It is an excellent solution for industrial processes simulation, including a module of human simulation which allows performing very accurate simulation of manual tasks. Moreover, several codes able to evaluate ergonomics indexes are already implemented and it is also allowed creating customized routines to implement other codes, plugins or interfaces. In this regard, a custom plugin has been developed in Visual C# to load the collected data. The plugin allows reading data stored on a CSV file, transferring them to the DHM and, hence, creating compound operations, according to experimental data, by using the function “CreateHumanCompoundOperation”.

The pseudocode for data transferring can be found in Fera et al. [4].

For this study, data have been re-sampled in order to reduce the number of micro-operations and, hence, the computational cost. Figure 6 shows the interpolating curve of the sampled data related to the posture angles trends, shown in Figure 5.

Motion data related to all the 60 investigated working cycles have been transferred and, hence, all the working cycles have been simulated.

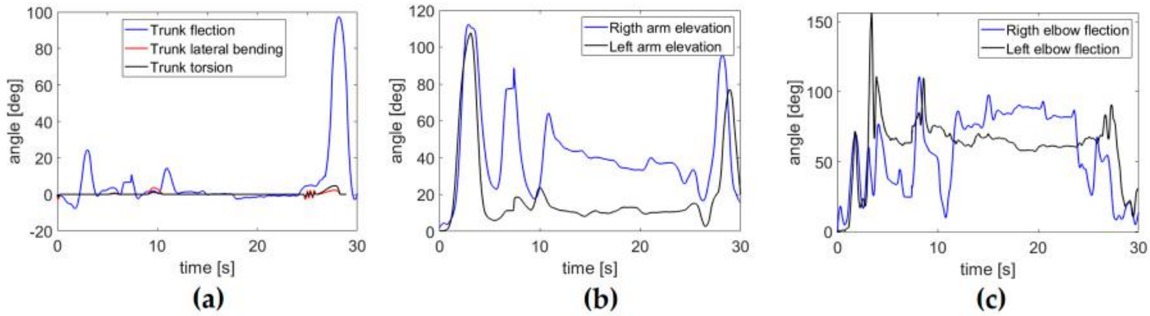


Figure 6. Posture angles smooth trends over one working cycle for: trunk (**a**), arms (**b**) and elbows (**c**).

3.3. Ergonomic Assessment

At the end of the simulation, numerical data are analyzed to perform the ergonomic assessment. As stated in Section 2.3, the main sources of injury risks are postures assumed by workers during each cycle of the work shift, exerted forces, MMH and repetitive actions; thus, each of this class has to be investigated by means of different ergonomic assessment methods.

In this case study, the methods chosen to evaluate each ergonomic category are the most feasible considering the type of working activity performed. They are listed below, with a short description:

- Working Postures: OWAS (Ovako Working Analysis System) method. It allows evaluating the whole body postures [37]. The method consists in the analysis of the posture to which a risk class is assigned, based on the values of postural angles. There are four classes (no risk, low risk, medium risk and high risk respectively) The index (I) for the whole cycle is evaluated according to the frequency (a, b, c, d) with which the four risk classes are found by means of the following equation:

$$I = [(a \cdot 1)+(b \cdot 2)+(c \cdot 3)+(d \cdot 4)] \cdot 100 \quad (2)$$

- Manual Material Handling: NIOSH (National Institute for Occupational Safety and Health) lifting equation, which is useful when manual lifting of loads is carried out [38]. The method is applied only if the weight of the handled object is higher than 3 kg. It assigns a score based on the vertical and horizontal displacements with which the object is handled, as well as the frequency of action and the type of grip. The Lifting Index (LI) is given by the ratio between the Loaded Weight (LW) and the Recommended Weight Limit (RWL):

$$LI = LW/RWL \quad (3)$$

where RWL depends on: worker's genre and age, vertical displacement of the object, maximum horizontal distance between the object and the body, angular dislocation of the object with respect to the sagittal plane, grip mode and lifting frequency;
- Repetitive actions: OCRA (Occupational Repetitive Actions) checklist, which is used to evaluate the biomechanical overload of upper limbs related to repetitive actions [39]. The checklist assigns a score based on the number of actions, fatigue (Borg scale), incongruous shoulder and wrist postures, type of grip, work organization, etc.;
- Force: in order to evaluate forces, a Force Solver based on Snook and Ciriello tables [35] was used. The tool provides the maximum force that can be exerted with respect to the posture assumed and the direction of application.

Table 2 reports the risk areas for the selected indexes.

Table 2. Risk areas for the selected indexes.

Index	Range Values		
	Low Risk	Medium Risk	High Risk
OWAS	<200	201–300	301–400
NIOSH	≤0.85	0.86–0.99	≥1
OCRA checklist	≤11	11.1–22.5	>22.5

It is worth noting that Equation (1), needed to perform the control phase of ergonomic indexes, is not applicable to all the chosen methods. In fact, Equation (1) is applied only when the index has a variability during the investigated cycles.

In this application, OWAS and OCRA indexes present variabilities along the working cycles, since fatigue, distractions or other factors may principally affect the working postures assumed by the worker; in contrast, exerted forces and NIOSH indexes remain the same in different cycles, since they depend on the positions of objects and on their weights, which are not variable.

3.3.1. Indexes Evaluation

In this section, the indexes evaluation is presented. The indexes have been evaluated by means of tools implemented in Tecnomatix Process Simulate, except for the OCRA checklist which has been filled-in by using data provided by the simulation.

Figure 7 shows the results concerning the OWAS index evaluation.

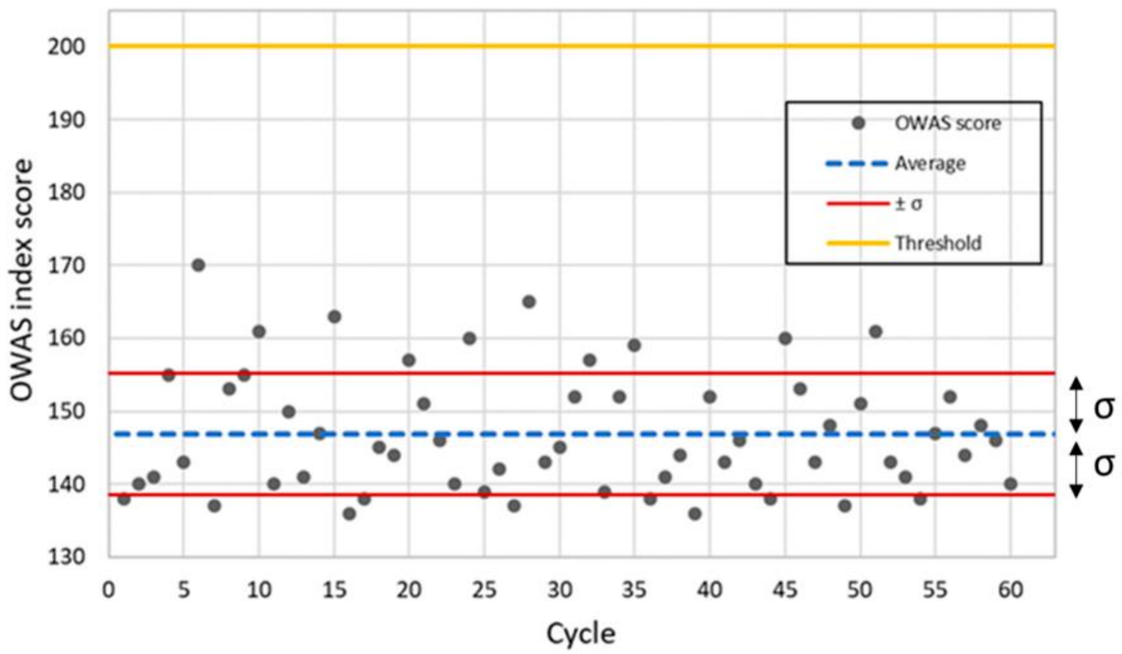


Figure 7. OWAS index scores for each working cycle and their statistical values.

The blue dotted line represents the average value of the OWAS index (I_{μ}), while red lines represent the standard deviation $\pm\sigma$ (I_{σ}) values. Table 3 resumes the statistical values and their comparison with medium risk threshold value (I_t).

Table 3. OWAS index score: statistical values.

	Average (I_{μ})	Standard Deviation (I_{σ})	Medium Risk Threshold Value (I_t)
OWAS index	146.9	8.3	200.0

By applying the Equation (1), it is possible to deduce that, as well as demonstrated by the results in Figure 6, it is widely verified:

$$I_{\sigma} = 8.3 < I_t - I_{\mu} = 53.1 \quad (4)$$

Thus, no further investigations are needed about working postures.

Regarding exerted forces, there are 4 operations that require the application of force. They are related to screwing operations (OP40 in Figure 3): the worker applies a counter-reaction force at the tightening end. The value of the force is the same for each screwing and in each working cycle, so there is no variability. It depends on the tightening torque (T), which is 30 Nm, and the lever arm (a), which for the used gun screwdriver (Figure 8b) is 100 mm. So the exerted force (F_{EX}) is given by the following equation:

$$F_{EX} = T/a = 300 \text{ N} \quad (5)$$

The Force Solver tool in Tecnomatix Process Simulate enables to evaluate the maximum applicable force, based on the posture and on the direction of application.

Since it is not possible to predict the direction of the tightening end force, the worst-case scenario has been considered, which corresponds to the direction, in the transverse plan, along which the minimum value of the maximum applicable force is depicted (Figure 8a). The maximum applicable force is equal to 44 N.

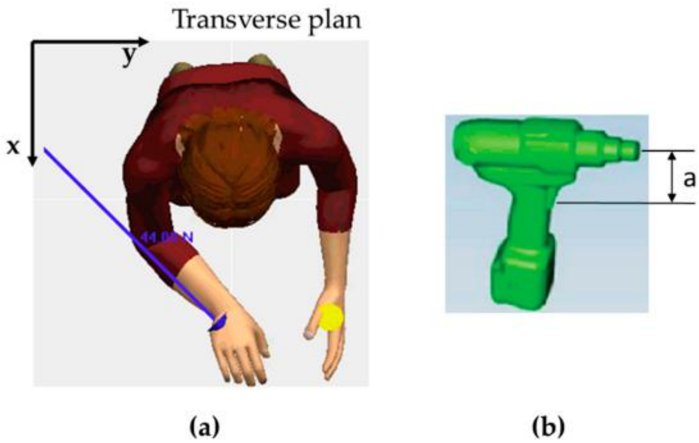


Figure 8. (**a**) Maximum applicable force. (**b**) Screwdriver lever arm.

Hence, the exerted force (F_{EX}) is widely higher than the maximum applicable force, so an improving solution is necessary.

Concerning MMH, NIOSH lifting index (LI) has been evaluated by considering two lifting operations that are performed for each task:

1. The first lifting regards the component 1, which has a weight of 6 kg. The heights of initial and final positions are 1400 mm and 900 mm respectively and the vertical displacement represents the main contribute to *RWL*. This component is handled with a frequency of 1 per cycle, thus during the work shift, a total of 900 components are lifted;
2. The second lifting concerns the assembly, which has a weight of 8 kg. The heights of initial and final positions are 900 mm 500 mm respectively. In this case, the main contribution to *RWL* is given by the maximum distance between the object and the body, which is equal to 400 mm. The frequency of handling is 1 per cycle too.

The NIOSH Lifting Index (LI) is constant throughout the working cycles and it is equal to:

$$LI = 0.96 \quad (6)$$

Concerning OCRA checklist, the scores varies along the cycles. Figure 9 shows the results concerning the right limb, which is the most stressed one.

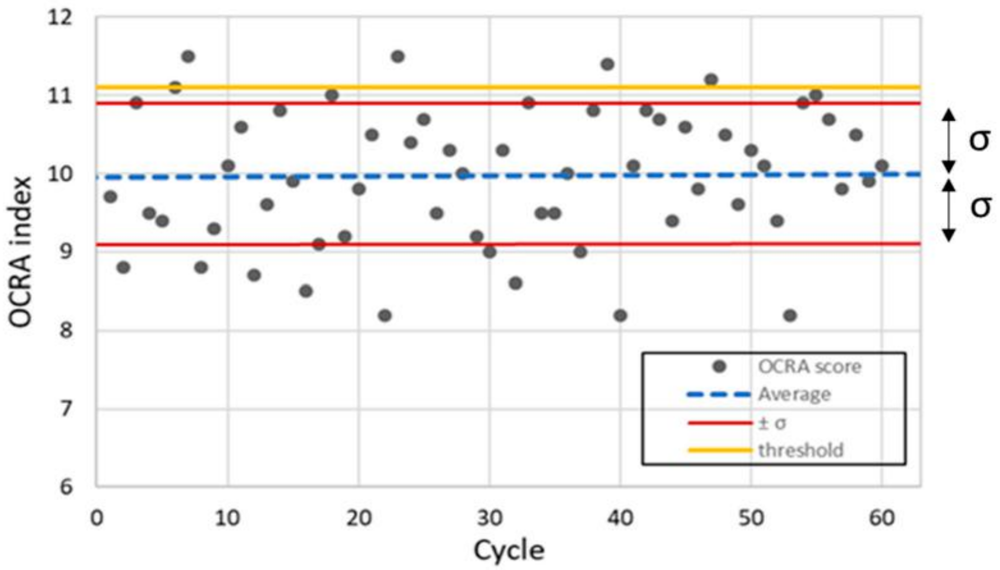


Figure 9. OCRA scores for each working cycle and their statistical values.

The blue dotted line represents the average value of the OCRA score (I_μ), while red lines represent the standard deviation (I_σ) values. Table 4 resumes the statistical values and their comparison with medium risk threshold value (I_t).

Table 4. OCRA score: statistical values.

	Average (I_μ)	Standard Deviation (I_σ)	Medium Risk Threshold Value (I_t)
OCRA index	10.0	0.9	11.0

By applying the Equation (1), it is possible to deduce that, as well as demonstrated by the results in Figure 9, it is verified:

$$I_\sigma = 0.9 < I_t - I_\mu = 1 \quad (7)$$

Thus, as well as for postures, also repetitive actions do not need further investigations, even if in this case the values are borderline between low risk and medium risk areas. This is deducible also from Figure 8, in which more than one cycle has OCRA score within medium risk area.

3.3.2. Critical Issues

Once the results about ergonomics indexes have been obtained, according to the methodological framework in Figure 1, if one or more indexes do not fall within the low risk area, it is appropriate to investigate which sub-phase or specific characteristic of the working cycle mostly contributes to the value of the index.

The analysis of the results related to the case study described shows that the values of exerted forces and the lifting index (NIOSH) exceed the threshold values for the low-risk area; hence, it is necessary an intervention for reducing the injury risk due to biomechanical overload.

Regarding the exerted forces, the value of the counter-reaction force, due to the tightening of the bolt, exceeds the maximum exercisable value. To decrease this force value, it is necessary to use a different type of screwdriver, with a higher lever arm. An angled screwdriver, with a distance between the spindle and the actuation button higher than a gun screwdriver, will significantly reduce the value of the counter-reaction force absorbed by the worker arm.

About the lifting index, according to NIOSH equation, it was deduced that the main contribution to is given by the vertical distance to be covered in handling and the maximum horizontal distance between the handled component and the body, which is excessive especially when the assembly is placed in the cart. In order to reduce the index, it is necessary to think about a reconfiguration of the workstation, lowering the shelf where the components are placed and raising the support surface of the assembly inside the cart.

The next section describes a possible modification of the workstation layout and the type of screwdriver. A simulation will show how this contributes to reduce the value of the risk indexes.

3.4. Proposal and Testing of Improving Solutions

According to the procedure shown in Figure 1, this section aims at proposing workstation layout and equipment changes in order to reduce the values of risk indexes.

As stated in the previous Section 3.3.2, NIOSH lifting index can be reduced by:

- reducing the vertical distance for the operations OP10 and OP20 of Figure 3, so the height of the shelf where the two components are located at the beginning of the working cycle;
- reducing the maximum horizontal distance between the assembly and the body for the operation OP60, which depends on the height of the support surface of the cart, where the assembly is placed.

For this purposes, the shelf has been modified: its height from the ground has been reduced from 1400 mm, as in the previous configuration, 1170 mm. Since the workbench is 900 mm above the ground, the vertical dislocation has been significantly reduced.

The new cart has been designed in order to significantly increase the height of the support surface from the ground: from 500 mm to 1030 mm. In addition, a break has been made to facilitate the positioning of the assembly. In this way the horizontal distance is reduced and the worker does not assume a posture with a large trunk flexion.

Figure 10 shows the new workstation layout.

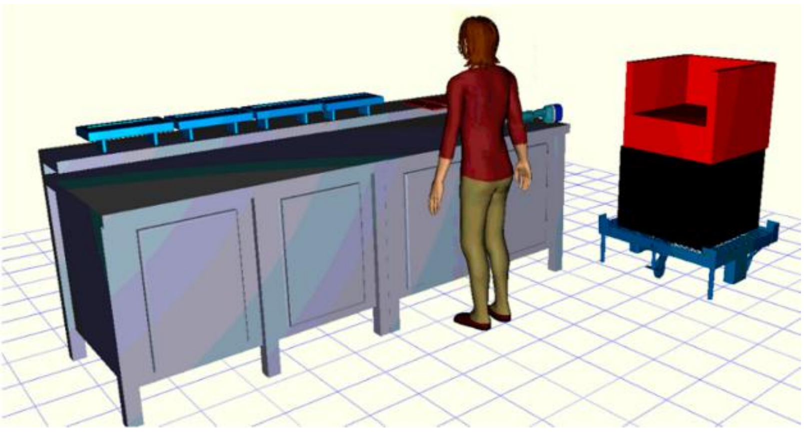


Figure 10. Workstation layout after equipment changes.

Concerning, the counter-reaction force due to the screwing operations, the gun screwdriver has been replaced with an angle screwdriver (Figure 11) with 300 mm lever arm (size "a"), very commonly found on the market.

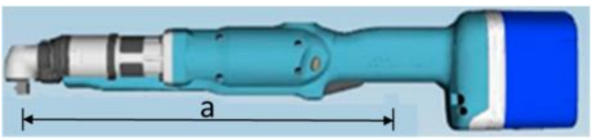


Figure 11. Angle screwdriver; size of lever arm equal to "a".

Figure 12 shows the working tasks after the workstation layout and equipment changes, simulated in Tecnomatix Process Simulate software environment.

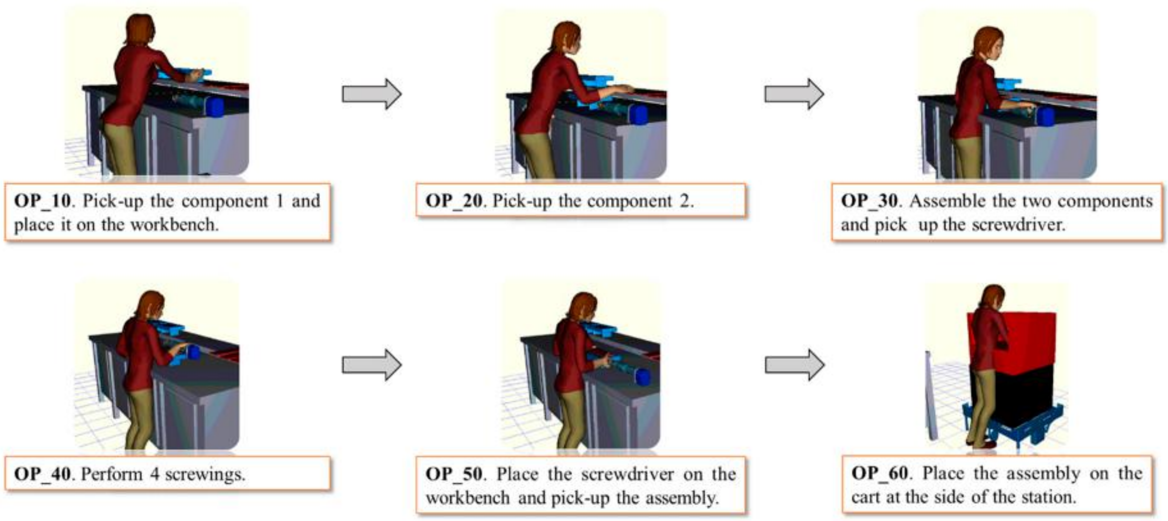


Figure 12. Working tasks after workstation layout and equipment changes.

A simulation has been run in order to numerically evaluate the 4 selected risk indexes. Table 5 shows the risk indexes values related to the new workstation configuration and the comparison with the risk index values evaluated in the previous configuration.

Table 5. Risk indexes comparison after workstation layout changes.

Index	Old Configuration	New Configuration	Reduction [%]
OWAS	146 (±8.3)	118	19.2
Exerted Force	300 N	100 N	66.6
NIOSH	0.96	0.59	38.5
OCRA checklist	10 (±0.9)	10	0

Results show that by modifying the workstation layout and equipment, the risk indexes drastically reduce.

Concerning the critical issues that emerged in the previous analysis, the NIOSH index has been reduced by 38.5% and now it falls within the low risk area. Concerning the exerted forces, despite a reduction of 66%, counter-reaction forces still exceed the upper limit. A further solution could be to change the characteristics of the bolted joints in such a way that a lower tightening torque is required.

However, a reduction in cycle time of about 4 s was also observed. Therefore, considering the unchanged production volume (900 pieces per shift), the recovery time for each cycle would increase. This could balance, at least in part, the biomechanical load due to the exerted forces, which exceed the maximum limit.

Finally, although it already fell within the low risk area, the OWAS index was also reduced by 10%.

4. Discussion

Ergonomic risk mapping of workstations is fundamental, as well as mandatory, for the companies to sustain high efficiency and commercial competitiveness without compromising workers' health and safety. However, the assessment of risk indexes is still carried out by means of observational techniques. This implies that their evaluation becomes high time consuming and, above all, affected by subjective considerations. This research was born with the aim of creating tools that can support experts in ergonomic screening and, above all, can ensure accuracy and repeatability of measurements. From this point of view, the DT can represent the best solution as it allows exploiting the computational capability by using numerical models in which real data, in this case related to human motion, are implemented.

Based on DT, the novel methodological framework here-in proposed has been developed to carefully investigate about ergonomics of manual working tasks in a manufacturing scenario. The case study has been introduced for proving its applicability and effectiveness.

As a consequence, the main implication for ergonomists is related to a significant reduction of the time for evaluating risk indexes, allowing them to focus mostly in identifying issues and proposing solutions. For data collection, we must consider the wearing and calibration time of the wearable system, that typically takes 10-15 min, which is acceptable in comparison to the very short time, in order of seconds, taken by the software for data analysis.

On the other hand, the proposed approach has some limitations, mainly due to the characteristics of software chosen for simulations. The analysis carried out in the present study has been performed by using Tecnomatix Process Simulate, one of the best solutions on the market. However, only some risk assessment methods codes are already implemented in the software (e.g., OCRA index code is not available) and this may require additional effort to implement the algorithm related to a specific index. Moreover, if a re-design of the work station is needed, for example if some issues related to the ergonomic screening tasks occur, the alternative solution is not automatically provided but it needs an analysis by an expert ergonomist. Lastly, in order to obtain a real time analysis, and so a proper DT, the motion capture system must be able to process and transfer data in real time. For the proposed case study, it has been used a system which requires off-line processing.

The literature about the use of human DT for the ergonomic evaluation of workstations is poor as already mentioned in the introduction. However, especially about the application of numerical models for evaluating ergonomic indexes, some comparison considerations can be reported.

First of all, the methodological framework is perfectly in line with the evolution of the topic in the era of industry 4.0, in the use of hardware and software tools for evaluating ergonomic indexes [40].

The present study offers a direct link with experimental data and the use of numerical model with respect to the studies by Caputo et al. [13,14], which are focused on the workstation design validation by considering ergonomic as a design parameter. The advantage offered by this novel methodological framework is the realization of a cyber-physical system in which the motion data of the worker are directly implemented in the simulation model. This offers the possibility to use a numerical model that, on the basis of experimental data, is able to provide ergonomists with a fast and reliable ergonomic screening tool for monitoring the real production.

Similarly, Bortolini et al. [41] carried out a complete ergonomic analysis by setting up an automatic procedure through an optical motion capture system. They used kinematic data for evaluating several risk indexes (OWAS, REBA, EAWS, etc.), demonstrating how such kind of procedures make the ergonomic assessment fast, reliable and objective. The here-in proposed methodology has several advantages compared to [41]. Firstly, the use of a wearable motion capture system allows collecting data directly in a working environment, during the normal production shift, which would be complicated by using optical devices that, although very accurate, are bulky and require a fine calibration. Moreover, the possibility to observe the simulation of the working activity allows immediately identifying possible anomalies and the simulation itself can be used as a training tool for other workers.

An analogous approach has been proposed in Grandi et al. [42], which through virtual environments proposes a workflow for ergonomic assessment during the design with the possibility of using immersive reality devices. However, also this procedure is not completely appropriate as not applicable in a factory environment, where, due to the tight spaces and the high focus that a work task requires, it is not easy to use immersive reality tools.

Zhang et al. [43] proposed a mathematical model for assessing ergonomics and optimizing the assembly line. Although it is a very interesting approach, the model considers only the OCRA index, evaluated according to traditional observational techniques.

The virtual simulation environment described in [17] by using an optical motion capture system is a good tool for the control of working time and quality in a manual manufacturing process. It could benefit from the use of the evaluation ergonomic indexes here proposed to improve the ergonomic performances in the manufacturing production scenario, even though a more performing motion capture device than Microsoft Kinect should be adopted.

The methodology proposed in the present is a contribution towards an innovative way for the ergonomic monitoring of manual workstations and it also contributes at reducing the literature gap about the topic. Collecting experimental data related to real workers, who perform their working tasks during normal working shifts, is extremely fundamental for a correct evaluation, especially in terms of objectivity and repeatability of measurements. In addition, by using cutting-edge technologies, it is possible to perform the ergonomic assessment in real time, with immediate feedback, allowing a significant reduction in evaluation time and, therefore, costs. This approach could be very beneficial for plant ergonomists and occupational physicians.

5. Conclusions

This paper presents a methodological framework, based on Digital Twin, aimed to assess the ergonomic performance in a manufacturing production scenario. Implementing motion data, collected during the working activities, in a virtual 3D scenario allows performing the ergonomic screening of the investigated workstation. In this way it is possible to evaluate the desired risk indexes and figure out eventual production issues.

A case study regarding a simple assembly task has been conducted in a laboratory environment to demonstrate the effectiveness of the proposed framework. Data related to working postures have been collected by a wearable inertial motion tracking system for 60 consecutive working cycles and transferred to the Digital Twin.

In this way it has been possible to easily evaluate four risk indexes related to working postures, exerted forces, material manual handling and repetitive actions, sources of biomechanical overload.

The first simulation figured out issues related to exerted forces and material manual handling, whose values overcame the upper limit of the low risk area. Hence, workstation layout and equipment changes have been proposed and a further time-based simulation has been run to test the solutions. The subsequently ergonomic assessment showed a significant reduction of the risk indexes.

It is fundamental to underline that the traditional ergonomic screenings, carried out in an observational way, are time consuming procedures which require several hours of work. The framework herein proposed allows drastically reducing the evaluation times as well as making the assessment objective and repeatable.

In summary, three important key elements can be pointed out in the present study:

- it has been demonstrated the possibility to assess ergonomics through an automated approach by implementing experimental data, collected by wearable sensors, in a simulation environment, creating the DT of a real manual workstation;
- this approach allows evaluating ergonomics in a faster and more accurate way than manual analysis. In fact, the subjective judgments of analysts are avoided, and the assessment becomes objective based on the used ergonomic indexes;
- the workstations can be continuously monitored, assessing the ergonomic indexes whenever necessary (e.g., in case of change of production volume). Moreover, ergonomists and engineers can identify critical situations, based on real data, and proposing solutions (such as the workstation layout change) to reduce risk indexes. Hence, new solutions may be verified by means of numerical simulation before their implementation.

The procedure described in this paper offers an improvement over current ergonomic screening techniques, but a bigger advantage will come from the use of devices capable of transferring data in real time, providing immediate ergonomic analyses.

Author Contributions: Conceptualization, A.G., M.C., M.F., S.G.; methodology, A.G., M.C.; software, A.G., M.C.; formal analysis, A.G., M.C., M.F., S.G.; investigation, A.G., M.C.; data curation, A.G., M.C.; writing—original draft preparation, A.G., M.C., S.G.; writing—review and editing, M.F., S.G.; visualization, A.G., M.C.; supervision, M.F., S.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: This work was supported by the University of Campania *Luigi Vanvitelli* under SCISSOR Project—V:alere program 2019.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Alessio, L.; Franco, G.; Tomei, F. *Trattato di Medicina del Lavoro*; Piccin: Padua, Italy, 2015.
2. Bernard, B.; Putz-Anderson, V. *Musculoskeletal Disorders and Workplace Factors; a Critical Review of Epidemiologic Evidence for Work-Related Musculoskeletal Disorders of the Neck, Upper Extremity and Low Back*; NIOSH Publications: Cincinati, OH, USA, 1997.
3. Takala, E.P.; Pehkonen, I.; Forsman, M.; Hansson, G.A.; Mathiassen, S.E.; Neuman, W.P.; Sjøgaard, G.; Veiersted, K.B.; Westgaard, R.H.; Winkel, J. Systematic Evaluation of Observational Methods Assessing Biomechanical Exposures at Work. *Scand. J. Work Environ. Health* **2010**, *36*, 3–24. [CrossRef]
4. Fera, M.; Greco, A.; Caterino, M.; Gerbino, S.; Caputo, F.; Macchiaroli, R.; D'Amato, E. Towards Digital Twin Implementation for Assessing Production Line Performance and Balancing. *Sensors* **2020**, *20*, 97. [CrossRef]
5. Cao, X.; Cheng, P.; Chen, J.; Sun, Y. An Online Optimization Approach for Control. *IEEE Trans. Ind. Inform.* **2013**, *9*, 439–450. [CrossRef]
6. Kadir, B.A.; Broberg, O.; da Conceição, C.S. Current research and future perspectives on human factors and ergonomics in Industry 4.0. *Comput. Ind. Eng.* **2019**, *137*, 106004. [CrossRef]
7. Atzori, L.; Iera, A.; Morabito, G. The Internet of Things: A survey. *Comput. Netw.* **2010**, *54*, 2787–2805. [CrossRef]
8. Kuruganti, U. 22—Sensors for Monitoring Workplace Health. In *Bioelectronics and Medical Devices*; Woodhead Publishing Series in Electronic and Optical Materials; Woodhead Publishing: Cambridge, UK, 2019; pp. 537–553.
9. Lim, S.; D'Souza, C. A narrative review on contemporary and emerging uses of inertial sensing in occupational ergonomics. *Int. J. Ind. Ergon.* **2020**, *76*, 102937. [CrossRef]
10. Hajifar, S.; Sun, H.; Megahed, F.M.; Jones-Farmer, L.A.; Rashedi, E.; Cavuoto, L.A. A forecasting framework for predicting perceived fatigue: Using time series methods to forecast ratings of perceived exertion with features from wearable sensors. *Appl. Ergon.* **2021**, *90*, 103262. [CrossRef]
11. Banks, J.; Carson, J.S., II; Nelson, B.L.; Nicol, D.M. *Discrete-Event System Simulation*, 5th ed.; Essex: Harlow, UK, 2019.
12. Alkan, B.; Vera, D.; Ahmad, M.; Ahmad, B.; Harrison, R. A model for complexity assessment in manual assembly operations through predetermined motion time systems. *Proc. CIRP* **2016**, *44*, 429–434. [CrossRef]
13. Caputo, F.; Greco, A.; Fera, M.; Macchiaroli, R. Digital twin to enhance the integration of ergonomics in the workplace design. *Int. J. Ind. Ergon.* **2019**, *71*, 20–31. [CrossRef]
14. Caputo, F.; Greco, A.; Fera, M.; Macchiaroli, R. Workplace design ergonomic validation based on multiple human factors assessment methods and simulation. *Prod. Manuf. Res.* **2019**, *7*, 195–222. [CrossRef]
15. Makarova, I.; Khabibullin, R.; Belyaev, E.; Mavrin, V.; Verkin, E. Creating a safe workin environment via analyzing the ergonomic parameters of workplaces on an assembly conveyor. In Proceedings of the 2015 International Conference on Industrial Engineering and Systems Management (IESM), Seville, Spain, 21–23 October 2015. [CrossRef]

16. Case, K.; Hussain, A.; Marshall, R.; Summerskill, S.; Gyi, D. Digital human modelling and the ageing workforce. *Procedia Manuf.* **2015**, *10*, 3694–3701. [CrossRef]
17. Tarallo, A.; Mozzillo, R.; Di Gironimo, G.; De Amicis, R. A cyber-physical system for production monitoring of manual manufacturing processes. *Int. J. Interact. Des. Manuf.* **2018**, *12*, 1235–1241. [CrossRef]
18. Sanjog, J.; Patel, T.; Karmakar, S. Occupational ergonomics research and applied contextual design implementation for an industrial shop-floor workstation. *Int. J. Ind. Ergon.* **2018**, *72*, 188–198. [CrossRef]
19. Catarci, T.; Firmani, D.; Leotta, F.; Mandreoli, F.; Mercella, M.; Sapio, F. A conceptual architecture and model for smart manufacturing relying on service-based digital twin. In Proceedings of the 2019 IEEE International Conference on Web Services (ICWS), Milan, Italy, 8–13 July 2019.
20. Malik, A.A.; Bilberg, A. Digital twins of human robot collaboration in a production setting. *Procedia Manuf.* **2018**, *17*, 278–285. [CrossRef]
21. Li, C.; Fahmy, A.; Sienz, J. An augmented reality based human-robot interaction interface using Kalman filter sensor fusion. *Sensors* **2019**, *19*, 4586. [CrossRef]
22. Nikolakis, N.; Alexopoulos, K.; Xanthakis, E.; Chryssolouris, G. The digital twin implementation for linking the virtual representation of human-based production tasks to their physical counterpart in the factory floor. *Int. J. Comput. Integr. Manuf.* **2019**, *31*, 1–12. [CrossRef]
23. Zheng, Y.; Yang, S.; Cheng, H. An application framework of digital twin and its case study. *J. Ambient Intell. Humaniz. Comput.* **2019**, *10*, 1141–1153. [CrossRef]
24. Ma, J.; Chen, H.; Zhang, Y.; Guo, H.; Ren, Y.; Mo, R.; Liu, L. A digital twin-driven production management system for production workshop. *Int. J. Adv. Manuf. Technol.* **2020**, *110*, 1385–1397. [CrossRef]
25. Kunath, M.; Winkler, H. Integrating the Digital Twin of the manufacturing system into a decision support system for improving the order management process. *Procedia CIRP* **2018**, *72*, 225–231. [CrossRef]
26. Harvard, V.; Jeanne, B.; Lacomblez, M.; Baudry, D. Digital twin and virtual reality: A co-simulation environment for design and assessment of industrial workstation. *Prod. Manuf. Res.* **2019**, *7*, 472–489.
27. Caputo, F.; Greco, A.; D'Amato, E.; Notaro, I.; Lo Sardo, M.; Ghibaudo, L. A Human Posture Inertial Tracking System for Ergonomic Assessment. *Adv. Intell. Syst.* **2019**, *825*, 173–184.
28. Giacomazzi, F.; Conti, C. *Manuale di Gestione Della Produzione*, 3rd ed.; ISEDI: Milano, Italy, 1975.
29. Tecnomatix. Available online: https://www.plm.automation.siemens.com/global/en/products/tecnomatix/ (accessed on 28 July 2020).
30. Delmia. Available online: https://www.3ds.com/products-services/delmia/ (accessed on 28 July 2020).
31. *ISO 11226: Ergonomics—Evaluation of Static Working Postures*; ISO: Geneva, Switzerland, 2000.
32. *ISO 11228-1: Ergonomics—Manual Handling—Part 1: Lifting and Carrying*; ISO: Geneva, Switzerland, 2006.
33. *ISO 11228-2: Ergonomics—Manual Handling—Part 2: Pushing and Pulling*; ISO: Geneva, Switzerland, 2006.
34. *ISO 11228-3: Ergonomics—Manual Handling—Part 3: Handling of Low Loads at High Frequency*; ISO: Geneva, Switzerland, 2007.
35. Snook, S.H.; Ciriello, V.M. The design of manual handling tasks: Revised tables of maximum acceptable weight and forces. *Ergonomics* **1991**, *34*, 1197–1213. [CrossRef]
36. Masali, M. *L'Italia si Misura. Vademecum Antropometrico per il Design e L'ergonomia. Vent'anni di Ricerche (1990–2010)*; Aracne Editore: Rome, Italy, 2015.
37. Louhevaara, V.; Suurnäkki, T.; Hinkkanen, S.; Helminen, P. *OWAS: A Method for the Evaluation of Postural Load during Work*; Institute of Occupational Health, Centre for Occupational Safety: Helsinki, Finland, 1992.
38. Waters, T.R.; Putz-Anderson, V.; Garg, A. *Applications Manual for the Revised NIOSH Lifting Equation*; National Institute for Occupational Safety and Health: Washington, DC, USA, 1994.
39. Colombini, D.; Occhipinti, E.; Alverez-Casado, E. *The Revised OCRA Checklist Method*; Editorial Factor Humans: Barcelona, Spain, 2013.
40. Mgbemena, C.E.; Tiwari, A.; Xu, Y.; Prabhu, V.; Hutabarat, W. Ergonomic evaluation on the manufacturing shop floor: A review of hardware and software technologies. *CIRP J. Manuf. Sci. Technol.* **2020**, *30*, 68–78. [CrossRef]
41. Bortolini, M.; Gamberi, M.; Pilati, F.; Regattieri, A. Automatic assessment of ergonomic risk for manual manufacturing and assembly activities through optical motion capture technology. *Proc. CIRP* **2018**, *72*, 81–86. [CrossRef]

42. Grandi, F.; Peruzzini, M.; Zanni, L.; Pellicciari, M. An automatic procedure based on virtual ergonomic analysis to promote human-centric manufacturing. *Proc. Manuf.* **2019**, *38*, 488–496.
43. Zhang, Z.; Tang, Q.H.; Ruiz, R.; Zhang, L. Ergonomic risk and cycle time minimization for the U-shaped worker assignment assembly line balancing problem: A multi-objective approach. *Comput. Oper. Res.* **2020**, *118*, 104905. [CrossRef]

applied sciences

MDPI

Article

Towards Integrated Digital Twins for Industrial Products: Case Study on an Overhead Crane

Juuso Autiosalo [1,*], Riku Ala-Laurinaho [1], Joel Mattila [1], Miika Valtonen [2], Valtteri Peltoranta [3] and Kari Tammi [1]

1 Department of Mechanical Engineering, Aalto University, 02150 Espoo, Finland; riku.ala-laurinaho@aalto.fi (R.A.-L.); joel.mattila@aalto.fi (J.M.); kari.tammi@aalto.fi (K.T.)
2 Remion Ltd, 33210 Tampere, Finland; miika.valtonen@remion.com
3 Konecranes Global Corporation, 05830 Hyvinkää, Finland; valtteri.peltoranta@konecranes.com
* Correspondence: juuso.autiosalo@aalto.fi

Featured Application: The initial version of a digital twin for an overhead crane with the main focus on providing data for machine designers and maintainers to support decision making and additional features for operation.

Abstract: Industrial Internet of Things practitioners are adopting the concept of digital twins at an accelerating pace. The features of digital twins range from simulation and analysis to real-time sensor data and system integration. Implementation examples of modeling-oriented twins are becoming commonplace in academic literature, but information management-focused twins that combine multiple systems are scarce. This study presents, analyzes, and draws recommendations from building a multi-component digital twin as an industry-university collaboration project and related smaller works. The objective of the studied project was to create a prototype implementation of an industrial digital twin for an overhead crane called "Ilmatar", serving machine designers and maintainers in their daily tasks. Additionally, related cases focus on enhancing operation. This paper describes two tools, three frameworks, and eight proof-of-concept prototypes related to digital twin development. The experiences show that good-quality Application Programming Interfaces (APIs) are significant enablers for the development of digital twins. Hence, we recommend that traditional industrial companies start building their API portfolios. The experiences in digital twin application development led to the discovery of a novel API-based business network framework that helps organize digital twin data supply chains.

Keywords: digital twins; crane; machine design; integration; maintenance; operation; API; open source

Citation: Autiosalo, J.; Ala-Laurinaho, R.; Mattila, J.; Valtonen, M.; Peltoranta, V.; Tammi, K. Towards Integrated Digital Twins for Industrial Products: Case Study on an Overhead Crane. *Appl. Sci.* **2021**, *11*, 683. https://dx.doi.org/10.3390/app11020683

Received: 23 October 2020
Accepted: 21 December 2020
Published: 12 January 2021

1. Introduction

Digital twin (DT) represents a new paradigm for the Industrial Internet of Things. DTs are linked to a real-world counterpart and leverage several technologies and other paradigms, such as simulation, artificial intelligence, and augmented reality, for optimizing the operation of the counterparts. DTs are being built at an accelerating pace in both industry and academia and the digital twin term has reached multiple expressions of recognition, such as being among the IEEE Computer Society's Top 12 Technology Trends for 2020 [1]. The roots of the digital twin concept are in mirroring physical systems as exemplified by two seminal publications: Grieves and Vickers [2] described digital twins as a set of virtual information describing a potential or actual physical manufactured product and NASA [3] described DT as an integrated ultra-realistic simulation that combines physical models and sensor data. The purpose and use cases of the concept have since been further described in further leading publications [4–6]. These concentrate on various areas of mechanical engineering, which is a trend continued by the majority of digital twin applications as shown by multiple review articles [7–12]. More recently, other domains, such as healthcare [13], business [14], and buildings [15], are starting to adopt digital twins as well. For a more in-depth exploration of the background of the digital twin concept, we refer to Section II in Autiosalo et al. [16].

Several earlier studies make recommendations for digital twin research. There seems to be a lot to be improved, as Liu et al. [11] conclude their review by stating that current digital twin literature cannot be inherited by other researchers. Other studies point out specific issues, including both technical aspects as well as considering humans as crucial actors in twin development. Marmolejo-Saucedo et al. [17] list the integration of information technology and the integration of partner companies as research issues. Barricelli et al. [10] list the cost of development and human interaction as challenges and call for a sociotechnical and collaborative approach in designing digital twins. Holler et al. [18] identify three topics for future research agenda: information architectures and models, specific applications, and the role of the human. Tao and Qi [19] recognize the need for experts from several disciplines and the need to make building digital twins easier. They also state that there should be physical 'innovation hubs' that are accessible to experts from different fields. Parmar et al. [14] recognize people and their skills as an essential factor in adopting digital twins. We condense these research issues into the following research question: "How to build integrated digital twins for industrial products?" This study aims to answer this question through a case study on the digital twin development journey of an industrial overhead crane.

Despite the high amount of manufacturing-related DT publications, only few DTs have been developed specifically for cranes. Moslått et al. [20] developed and validated a simulation-and-control focused digital twin for an offshore crane for lift planning purposes. Moi et al. [21] experimentally verified a real-time simulation of strain on a small-scale boom crane to prove that simulation-based virtual sensors can be used for condition monitoring. Szpytko and Duarte [22] developed a statistical decision-making model to efficiently schedule maintenance breaks for port gantry cranes. Additionally, an overhead crane is mentioned as a part of an ontology focused roll grinding simulation configurator [23].

The main research data of this study comes from the digital twin development journey of an overhead crane "Ilmatar" [24]. Most of the development was implemented as an industry-university research project called "DigiTwin" which is referred to as "the project" in this study. Many parts of the project have been published earlier as separate works, and this study combines these components together to form the whole digital twin of the crane as depicted in Figure 1 and gathers together the development experiences to draw practical DT development recommendations. Hence, the key contributions of the paper are:

- Presenting a multi-component digital twin of an overhead crane built by industry and university representatives. (Figure 1 and Section 3)
- Describing lessons learned on how to develop integrated digital twins. (Focused observations in Section 4.4, further material in the whole Section 4.)
- Identifying easy-to-use Application Programming Interfaces (APIs) as a significant practical enabler and requirement for creating integrated digital twin applications. Providing guidelines on how to leverage APIs efficiently. (Sections 3.15, 4.3, and 4.4.)

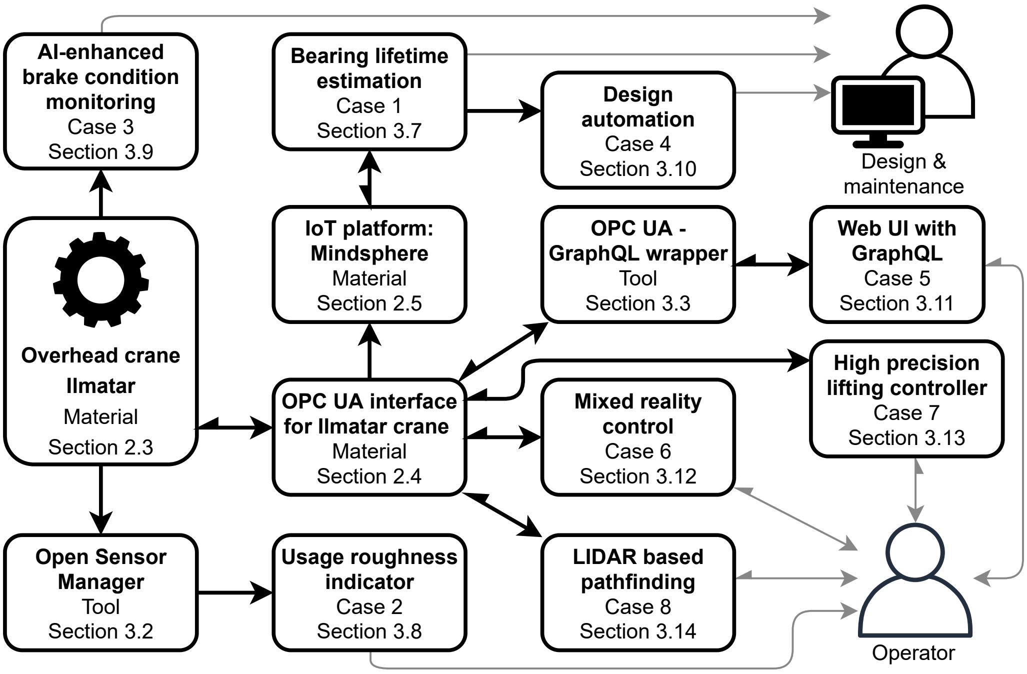


Figure 1. Overview of the digital twin components developed for the overhead crane. The black arrows depict data flow through a technical Application Programming Interface (API) and gray arrows are human-machine interfaces. The full arrowhead depicts primary information flow and half arrowhead secondary information flow, such as control. Each component is further described in the mentioned section and the APIs are further described in Section 3.15.

The digital twin of the Ilmatar crane is not yet finished as most of its components are separate from each other from the user perspective. Ongoing research and development efforts concentrate on integrating these components together. Meanwhile, this study shares the experiences of one digital twin development journey, contributing its share to the overall body of knowledge.

1.1. Hypotheses

Eight hypotheses on digital twins and related matters were made during the preparation and initial phases of the project. The first six are common assumptions mentioned in the project plan, the seventh was generated previously as a result of a practical study [25] in the same physical crane environment, and the eighth was a commonly agreed development direction at the start of the project. The hypotheses are:

1. Digital twin transforms data from a physical product to useful knowledge. Digital twin offers data and knowledge to all stakeholders across the product lifecycle.
2. Digital twin integrates digital models and data from different sources and providers and offers a customized view for each stakeholder.
3. Digital twin enables networking and business.
4. Machine design dimensioning and product development processes can be redefined with the true usage and maintenance data provided by a digital twin.
5. The overhead crane located at university premises acts as an excellent development platform, offering industrially relevant applications, such as plugging digital twin as part of product configuration, design, and life-cycle management.
6. Acting as an interface for all Industrial Internet data is one of the most important functions of a digital twin, enabling the efficient use of a vast amount of data.
7. "Using existing APIs enables fast prototyping, and bringing 'developer culture' from the 'software world' to the 'physical world' enables faster prototyping/product development cycles" [25].

8. Digital twin can be built without selecting a central visualization and simulation model.

The hypotheses were used as guiding principles during the project rather than being taken under taken under specific examination as the project concentrated on building practical use cases. The hypotheses are included in this paper to represent priors and motivation of the work as well as to act as a tool for discussing the results of the project in Section 4.2.

2. Methods and Materials

We used two research methods for this study: Participation Action Research (presented in Section 2.1) to gather data from the industry–university collaboration project, and the basic principles of Grounded Theory (Section 2.2) to draw conclusions from the data. Main materials used for the study include the overhead crane Ilmatar (Section 2.3), its data interface (Section 2.4), and an IoT platform (Section 2.5).

2.1. Participatory Action Research

Participatory Action Research (PAR) [26] refers to a research data acquisition method in which the researchers themselves participate in the activity they are analyzing. PAR is a subcategory of action research and has been used in various fields of social science research for decades. PAR enables rich data collection but brings in the risk of researcher bias.

This study used PAR to acquire data from the development process of the digital twin of the Ilmatar crane. We focus our analysis on two types of qualitative data: technical data on the contents of the digital twin, and socially oriented data on how the development process of that digital twin was implemented. The technical data mainly serves as an example of what a digital twin of an industrial product can be, and the development process data are used to derive the steps that need to be taken when implementing a digital twin.

A significant portion of the digital twin development was made as a single collaborative industry–university research and development project called "DigiTwin". The project was implemented at Aalto University with four funded industrial partners and one supportive industrial partner that provided resources and attended meetings. Additionally, some activities were performed outside the project.

The data of the development journey were collected in the following ways: observation through participation, project meeting memos and other materials (e.g., internal presentation slides and emails), and presentation recordings of two seminar sessions. The data from the first two methods are confidential whereas the recordings are publicly available on Youtube (https://www.youtube.com/channel/UCJrkhYovV4V-PwqlJeW5bmQ/videos). Furthermore, a master's thesis [27] was conducted on the dynamics of university–business cooperation and is used as supportive material.

2.2. Grounded Theory

Grounded Theory [28] is a theory generation process that includes three main phases: (i) collect (qualitative) data, (ii) organize data into categories, and (iii) analyze the relations between these categories to generate theories. Ideally, each category has multiple data points and the relation between the majority of the data points between the two categories is the same.

In this study, we take a similar approach as Autiosalo et al. [16] in leveraging Grounded Theory, meaning we leverage the basic principles instead of strict coding practices. As a comparison to this approach, Josifovska et al. [29] implemented Grounded Theory with three distinct coding phases to identify main building blocks and their properties for digital twins to further create a digital twin reference framework.

Induction is used as a supportive theory generation method. When using induction, we observe a situation that causes an effect and afterward reason a rule that states that a certain situation leads to the observed effect. In this work, we observed causes and effects during the DT development project and could induce that easy-to-use (quick to learn and set up) programming interfaces are a requirement for efficient digital twin application development. The amount of data covered in this study is not yet enough to call this a theory.

2.3. Overhead Crane Ilmatar

Aalto University has a full-size industrial overhead crane called "Ilmatar" (Figure 2) installed in one of its laboratory spaces [24]. The crane is manufactured by Konecranes and has several smart features such as target positioning, sway control, load floating, and snag prevention. The lifting capacity of the crane is 3200 kg, installed movement area 9.0 m by 19.8 m, and maximum speed 32 m/min. The crane is connected to the Internet, and several statistics about the use of the crane are sent to the vendor cloud and a third party IoT platform. A subset of the crane Programmable Logic Controller (PLC) system is linked to an Open Platform Communications Unified Architecture (OPC UA) server which provides a two-way (read and control) interface to crane data. The crane does not perform regular production line tasks and therefore the crane usage profile differs from its industry counterparts, having lower usage activity and seldom high-load lifts.

Figure 2. Overhead crane Ilmatar at Aalto University Industrial Internet Campus.

The crane serves as a research, innovation, and education platform and has been used in collaboration with companies as well as part of student projects. The crane development environment was published as "Ilmatar OIE" (Open Innovation Environment) in November 2019 further described in Section 3.1. The environment is open by default with some of the resources publicly available and some after manually handled registration.

Ilmatar crane is a special case among cranes and industrial products for two main reasons. First, it sees an extraordinarily large number of research, development, and education activities and therefore creates a lot of unstructured data. This is highlighted by the amount of cases presented in Sections 3.7–3.14. Second, being located in a student laboratory of a university, it is primarily public by nature in terms of both physical access and data sharing. Experiences from publishing results from the crane environment are described in Section 3.16.

2.4. OPC UA Interface of the Crane

OPC UA is an industrial standard for communication continuously developed by OPC Foundation [30]. It defines information, message, communication, and conformance models to enable interoperability, is platform-independent, and supports several communication protocols and data formats [31]. OPC UA has become a common standard especially among PLC manufacturers and is an important enabler for Industrial Internet applications. It is used mainly inside local factory networks.

The OPC UA server of the crane allows monitoring its current status variables, such as position and speed. The crane can also be controlled via the interface. Description of the OPC UA interface is currently available in Ilmatar OIE via sign-up [32]. The server itself is not publicly on the Internet, but in a password protected network in the laboratory hall. Authentication is not currently required to access the OPC UA server, but to control the crane, a specific access code has to be written periodically to a certain node. Modifications to the OPC UA server, such as adding new nodes, are made by manufacturer technicians to ensure safety.

2.5. IoT Platform: MindSphere

IoT platform "MindSphere" by Siemens [33] has been used for data gathering from the crane since its installation. Data gathering is performed by a physical MindSpere-specific gateway "MindConnect Nano" [34] that reads data from the OPC UA server of the crane. MindSphere was also used for building the bearing lifetime estimation application shown in Section 3.7. MindSphere is built on top of open-source cloud platform "Cloud Foundry", getting its core capabilities from that, while the practical productization and usability solutions are MindSphere specific. The MindSphere instance used with the crane was changed from version 2 to version 3 during the project.

3. Results

This section presents the practical components of the digital twin and the essential related material. These include a documentation solution, tools, frameworks, and case descriptions related to the digital twin development journey of the Ilmatar crane. Most of the results have been described earlier in academic publications or project seminars, all of which are referred at the start of each section if applicable. This study adds any necessary details from the overall DT development perspective. The level of technical detail for the results is sparse, concentrating on providing a meta description for the basis of discussion on the development of (integrated) digital twins.

All results are purpose-specific assets built for the overhead crane Ilmatar. Some of the results were intentionally built as part of the digital twin, while others were separate development efforts connected to the crane. The criteria for including them in this study is that they could be included as part of the digital twin of Ilmatar in the future.

3.1. Documentation: Ilmatar Open Innovation Environment

Ilmatar OIE is a combination of digital and physical resources of the crane environment. A single public web page [32] acts as a start page, including a basic description of the environment and a list of resources. The page provides links to the rest of the published digital resources that are either in academic publications, GitHub, or in a web workspace "Eduuni" that requires registration. In addition, a lot of unpublished works have been developed in the environment, and some of those are mentioned on the web page. Physical resources are currently not very extensively documented, and guidance on those rather relies on personal on-location instruction. This is a fairly natural choice as the crane is a full-scale industrial device with the potential to make extensive damage to its environment even with its safety features. Each crane user has to complete safety training before operating the crane.

The environment balances on what to include as public resources based on three main factors: (1) administration effort, i.e., how much time can be used to publish and update the materials, (2) user experience, i.e., comprehensiveness and ease of access to the materials, and (3) safety, i.e., physical safety, cybersecurity, and protecting immaterial property. The balancing is currently made on-the-fly, allowing natural development mainly based on the number of users.

Ilmatar OIE is currently the most comprehensive public collection of information dedicated to the Ilmatar crane. It was not originally made to be a part of the digital twin, but the amount of meta-knowledge it contains makes it relevant also from the digital twin perspective; Ilmatar OIE partially fulfills the "Data link" feature described by Autiosalo et al. [16] by providing a collection of useful links to various resources of Ilmatar crane.

3.2. Tool: OSEMA

Open Sensor Manager (OSEMA) is an open-source web platform that enables the setup and modification of settings on microcontroller-based sensors. It was developed as a practical response to the multitude of existing IoT protocols and communication methods noticed during master's thesis work by Ala-Laurinaho [35]. The work was further developed and published in a journal article [36]. Currently, OSEMA has no publicly available instance, but users can install it to their own server with the source code available at GitHub [37].

OSEMA can be used for the no-code setup of sensors and therefore enables effortless retrofitting sensors to the Ilmatar crane, aiding in data collection. After sensor installation, the manager web page can be used to change the configuration of sensor nodes, including measurement settings and network parameters, remotely over the Internet. OSEMA offers a web user interface and a Representational State Transfer (REST) API for the monitoring and management of the sensor nodes. Ilmatar was equipped with OSEMA-managed distance sensors tracking the location of the bridge and trolley. In addition, a 3-axis accelerometer was installed on the hook of the crane. This sensor was used in the usage roughness indicator application presented in Section 3.8.

OSEMA allows data collection from the real-world entity, which is a crucial part of the digital twin concept. Therefore, the sensor manager is considered an enabler building block for the implementation of digital twins. OSEMA can be used as the "Coupling" feature of FDTF (described in Section 3.4), supporting the creation of integrated digital twins with its REST API.

3.3. Tool: OPC UA–GraphQL Wrapper

OPC UA–GraphQL wrapper is a server application that connects to an OPC UA interface to provide it as a GraphQL interface. The wrapper was developed as a master's thesis by Hietala [38] and presented and evaluated in a conference publication [39]. The source code is published as open source in GitHub [40]. A GraphQL wrapper was installed to the Ilmatar crane on a Raspberry Pi, and an example control application was made for it, presented in Section 3.11.

GraphQL is a query language and execution engine open-sourced by Facebook in 2015 [41]. Since then, GraphQL has become a popular query language among developers [42]. GraphQL overcomes multiple shortcomings of the popular REST API and is seen as a successor for them [43].

The OPC UA–GraphQL wrapper was developed after noticing a hindrance in the crane interface: OPC UA, even though being a common standard in the industry, is not familiar to web software developers and it is fairly complicated compared to, for example, REST APIs. The OPC UA–GraphQL wrapper makes data from the OPC UA server of the crane easier to access, therefore facilitating software development for the Ilmatar. The OPC UA–GraphQL wrapper also comes with a built-in node viewer, so a user can use any standard web browser to click through the nodes and their current values. As a downside, not all features (e.g., publish-subscribe) of OPC UA are yet supported.

The wrapper represents an important phenomenon in the development of integrated digital twins: adapters. It takes some of the development burdens away from application development. While OPC UA might be technically possible to implement in any project, it may not be feasible during projects with a limited time frame, for example, due to the lack of libraries for some programming languages. The wrapper is also located between two cultures: OPC UA servers are typically installed in closed intranet networks in factories, whereas GraphQL servers are typically available via the public Internet. The benefits of networked digital twins can only be achieved if the twins can send messages to each other, for which the public Internet currently seems the most prominent option. Hence, data adapters that couple operation data with the Internet are important enablers for digital twins, although security solutions still need to be developed before the public Internet can be used.

3.4. Conceptual Framework: Feature-Based Digital Twin Framework

The feature-based digital twin framework (FDTF) is a framework that aims to deepen the conceptual understanding of digital twins by identifying features of digital twins and combining them into building blocks of digital twins. FDTF was developed during Ilmatar digital twin development and published as a journal article by Autiosalo et al. [16] and presented in a project seminar [44].

The features and building blocks of digital twins are put into practice by software components. One software component usually fulfills more than one feature, exemplified by the observation that the features exist in different hierarchical levels. The suggested features are data link (which connects the features together), coupling, identifier, security (enablers), data storage, user interface, computation (resources), simulation model, analysis, and artificial intelligence (producers). The purpose of these features is to take the focus of conceptual digital twin development away from existing tools, and into the development of novel, digital twin-specific solutions. In a proposed process of digital twin development, you first define a business need-based use case, then select the features needed to accomplish the needs, and lastly implement a digital twin with software components that provide the desired features. Examining the building blocks of a digital twin from the feature perspective makes redundancies evident, helping clarify the mess created by software components that were not originally made for digital twin implementation.

The components are tied together with a data link, which is being developed towards a ready-made software tool in a follow-up project. There are also several partial implementations of this conceptual approach, such as the Message Queuing Telemetry Transport (MQTT) broker [45], the whole Semantic Web approach [46], API gateways by different providers, the digital twin definition language [47], and the DT Core presented in Section 3.5. These represent the practical work towards integrated digital twins, and the data link concept is a statement to combine these approaches to create more comprehensive digital twins.

Defining digital twins was seen as a necessary activity before starting digital twin development. FDTF was developed as a response to this need. It aims to help understand the nature of digital twins, so that they can be developed further as a concept, rather than staying limited to the old tools and ways of working. FDTF highlights the difference between visionary conceptual development and actual implementation with existing software.

3.5. Implementation Framework: DT Core

DT core is a modular building block for DT data processing. It was presented in a project seminar by Valtonen [48]. It is a step towards realization from FDTF to DT application implementation, such as the one described in Section 3.9. DT core represents the state of a twin and consists of data storage, logic, and interfaces.

The data storage contains state descriptive data, metadata, and management data. The descriptive data form the main information contents for the DT, consisting of "hard" data such as 3D models, component structures, and numerical IoT data. Metadata describes the meaning of those data and interpretation guidelines for the hard data, such as value ranges and units for the data. Management data provides higher level insight to the contents of the DT, such as prioritization if data comes from multiple sources, parameters for data processing, history of state changes, and also a prediction of future state from a historical viewpoint. The management data serves the logic side of the DT core.

The logic of the DT creates additional value by refining the DT data, application-level logic, and processing management. Data within the storage can be refined with methods such as inference trees and AI models, and key performance indicator calculations. Application-level logic serves use cases through automatic reprocessing of data and signal abnormality detection. Processing management keeps data up-to-date by sustaining a processing heartbeat and may include alerts that are modified according to user feedback.

The interfaces bring data in and out of the DT core. They consist of query and storage management and security policies. Several query protocols can be supported to apply CRUD (create, read, update, and delete) operations, stream and batch-based updates, and standard Structured Query Language (SQL) operations to the storage and analytics parts of DT core. Each query faces security policies and is authenticated when needed. By enabling data exchange between multiple external systems, the interfaces of the DT core can be used to create an active data ecosystem.

Layering is used as a way to manage the time frame of actions for different types of DTs. The physical twin operates in real-time and the layers are divided by the heartbeat of their update frequency. Data on the layers are updated in an IoT platform, asset DT, fleet DT, and strategic DT, which leverage 1 s, 1 min, 1 h, and 1 day heartbeats, respectively (Figure 3).

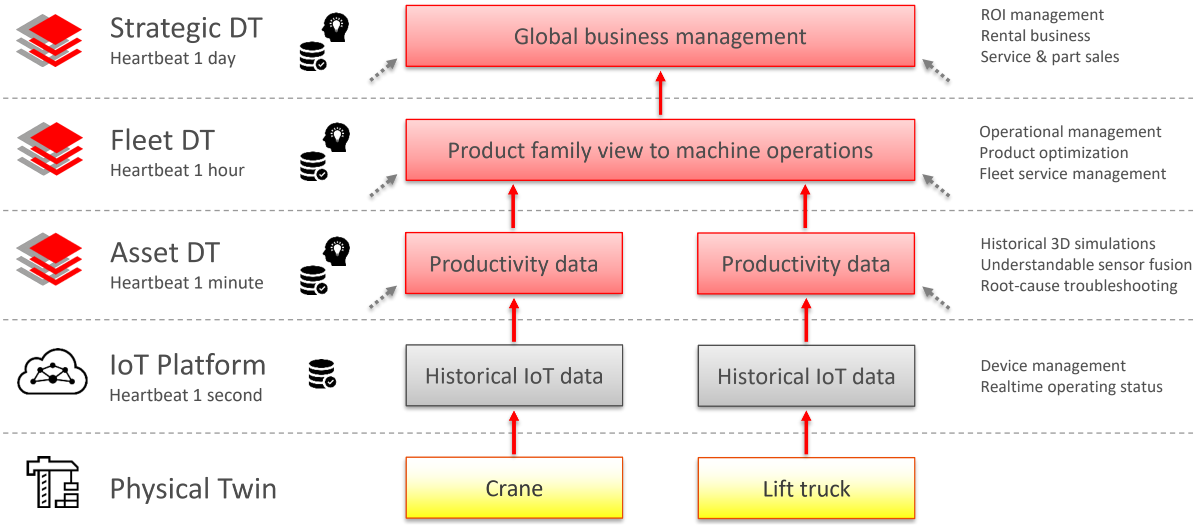


Figure 3. The layers of the DT core framework with update rate getting less frequent as data flows from physical twin to strategic decision-making level. Examples of use cases are given on the right.

DT core is currently an implementation framework for building digital twins that are focused on data analysis. It was used as a guiding principle in the development of the brake condition monitoring application presented in Section 3.9.

3.6. Information Framework: Digital Twin-Based Product Lifecycle Management

The digital twin-based product lifecycle management (DT-PLM) framework is a way to categorize the engineering content created from the initial idea to the product use phase. The framework was presented in a project seminar by Pantsar and Mäkinen [49]. The vision was used as a guiding principle during the project, although being an extensive framework, only selected parts of it could be implemented.

DT-PLM contains three categories: DT data, DT intelligence, and the real world, as shown in Figure 4. The DT data contain static representations of the product, whereas the active features in the intelligent part, such as simulation, bring digital twins to life, i.e., the DT becomes an active agent in cyberspace. The real world joins the lifecycle as the first prototype tests are performed and mirrored during the use phase of the product.

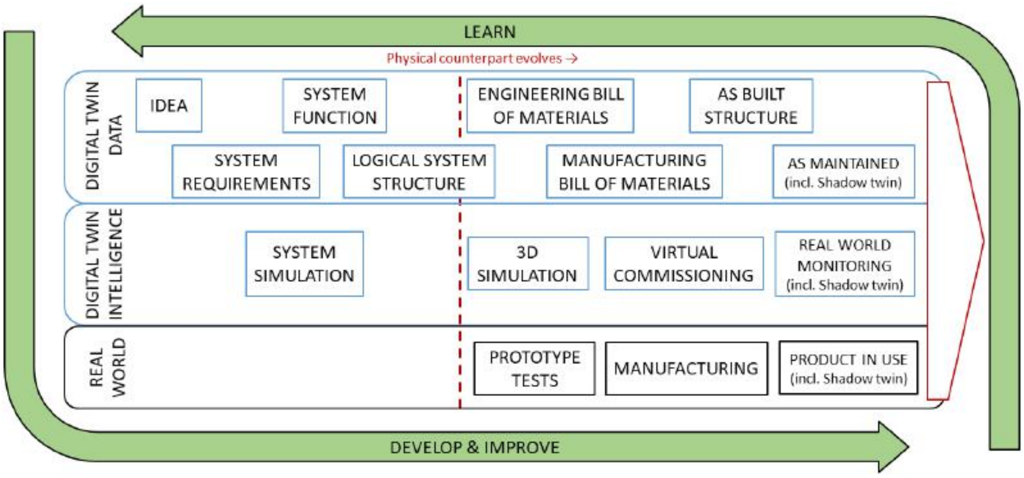


Figure 4. Digital twin-based product lifecycle management from product idea to product in use. The accumulating knowledge is used to develop and improve the product and to learn how to make future products better [49].

The original design data are referred to as "product DNA". Storing the DNA is the first step, the second is to use it in various tasks during the product lifecycle. For example, DNA can be used to predict problems by looking at the DNA shared by a product family. If the DNA is in a machine-readable format, it can be leveraged as an active part of the usage phase digital twin system. For example, a system simulation can be used to provide virtual sensor information when plugged into the real counterpart. As related work, the product DNA concept has also been used by Silvola [50] in similar meaning, although focusing especially on the uniqueness of DNA.

DT-PLM framework assists in understanding the practical information contents of digital twins from a product development perspective. This is an important perspective as a lot of potentially useful qualitative data are created during the product design phase. A challenge is that a lot of the design data are confidential and cannot be given even to buyers of the product. Currently, most of the design data are also prepared only for internal use, making them potentially unusable for others. A very strong business case would be needed for transferring the data to outsiders, but when the manufacturer is also the maintenance provider, using the design data in later phases of the lifecycle can be a gradual process.

3.7. Case 1: Bearing Lifetime Estimation

Bearing lifetime estimation case consists of a method and a proof-of-concept implementation for the closed-loop design of crane bearings by performing automated data analysis on crane usage data and bringing the analysis to a system used daily by product developers. The majority of the results described in this subsection were presented earlier in a project seminar [51] by Peltoranta and Autiosalo. A detailed description of the data analysis and visualization was given in a bachelor's thesis by Mattila [52].

The data flow for the case travels through several technical tools: (i) the physical crane with sensors, Programmable Logic Controller, and an OPC UA server, (ii) physical IoT gateway, (iii) cloud-based IoT platform with data storage, custom engineering formulas, and web technologies (e.g., JavaScript), and (iv) a PLM system. The crane produces the raw data and provides an OPC UA interface for the IoT gateway. The gateway sends selected usage data to an IoT platform that stores the data, performs customized analysis (kinematics-based virtual sensoring and bearing lifetime calculations) on the data, as well as hosts a visualization that is displayed on a PLM system. The data flow is shown in Figure 5.

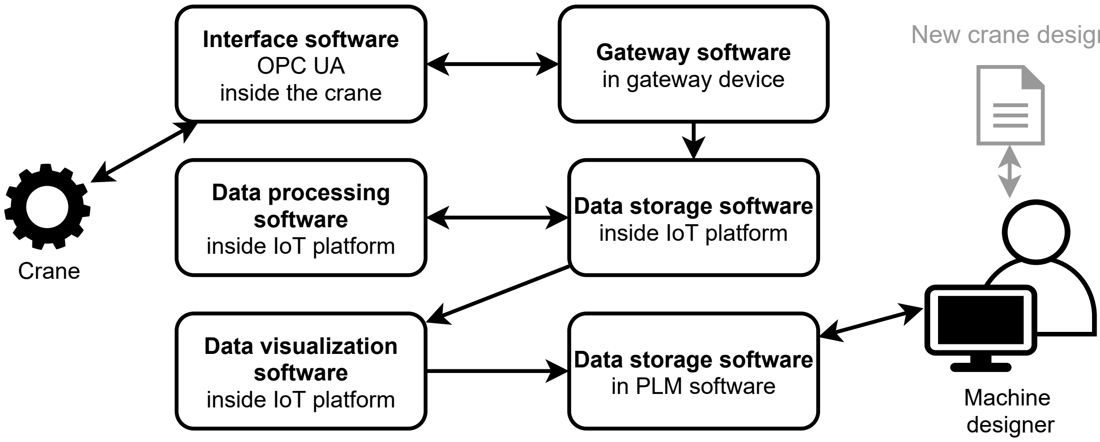
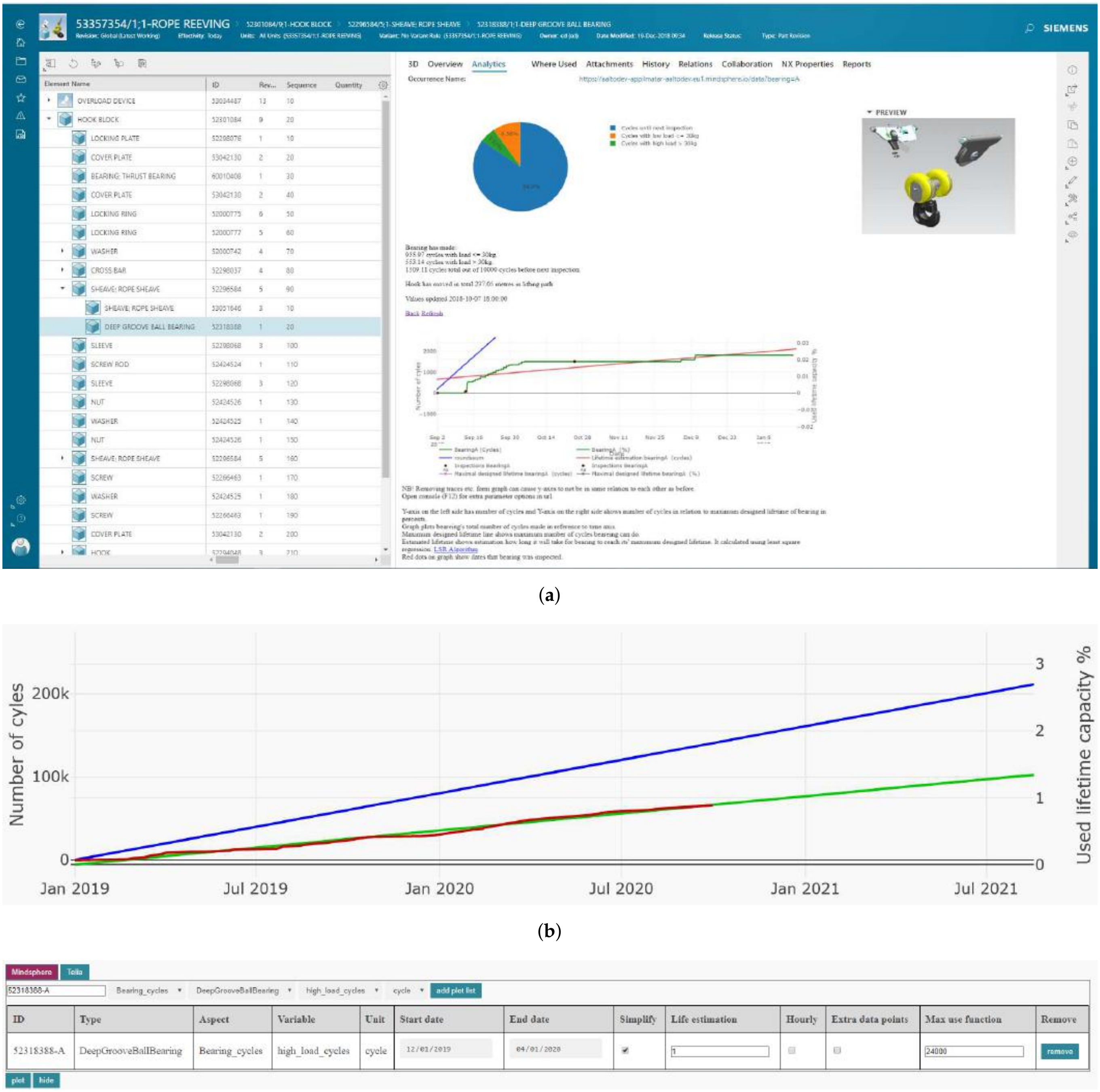


Figure 5. The data flow of the bearing lifetime estimation case. Designing a new crane is the purpose of providing the data, but not implemented during this study.

Collected raw data included the vertical position and the load of the crane, which were further refined with kinematic equations to determine the revolutions and load inflicted on each rope sheave bearing. The crane has three rope sheaves, each with a distinct number of revolutions. During the project, data collection for this implementation was straightforward with existing industrial tools, but presenting meaningful data in a meaningful way proved to be challenging.

Methodology for this case includes engineering design knowledge and practical data flow implementation tools from raw data generation to end-user visualization system. According to basic engineering dimensioning formulas, the lifetime estimations of ball bearings are based on the usage frequency and stress. When we get these data from the crane, we can compare the designed lifetime and the usage-based predicted lifetime of the component. For the example crane Ilmatar, the usage-based lifetime prediction for the bearings was approximately a thousand years due to the low usage frequency and light loading exerted on the crane used for research and education. With the lifetime prediction, a machine designer can determine if the component has been selected correctly. This knowledge can be leveraged to design new cranes that better fit their expected usage profile. Of course, the machine designer can perform estimations themselves, but the automated analysis saves their time for more demanding tasks and with large fleets, manual estimations may be infeasible. As future development, automated analysis can be used for triggering alerts if a crane is used more than expected.

The end-user interface of the case is the visualization that was shown inside a web-based PLM application (Figure 6a). Figure 6b shows the estimated usage (green line) of the bearing if the use of the crane continues similarly as during a selected time period (red curve). The time period and other parameters of the graph, such as the designed lifetime (pink line) that is used for drawing the reference line (blue line), can be changed either by changing parameters in the URL of the browser or with a configurator user interface shown in Figure 6c.



(a) (b) (c)

Figure 6. Screenshot samples of user interfaces for the bearing lifetime estimation case. (**a**) Product lifecycle management (PLM) software Teamcenter showing bearing usage data. The pie chart shows the types of usage and the graph shows timewise usage. The number of cycles is calculated from true usage data of the Ilmatar crane and the inspections are mockup for proof-of-concept visualization purposes. (**b**) A graph visualizing the past use (red curve), predicted expenditure according to the past usage (green), and the allowed linear usage to achieve the designed number of cycles (blue) in the designed lifetime of one of the bearings in the Ilmatar crane. The chosen lifetime for drawing the blue line is 100 years. (**c**) The user interface for selecting the information to be displayed on the graph.

The codebase for the web application was put into a GitLab instance upkept by university IT services and a continuous integration and continuous delivery (CI/CD) pipeline was made for the application using a Linux server provided by the university IT services. Any changes to the GitLab codebase, made with a git client or via a web page, triggered tests for the new application and if they succeeded, the changes were automatically deployed to the web application.

The case functions as a prototype to guide the creation of DTs for a larger fleet of cranes. The larger fleet supposedly enables further opportunities, such as usage profile categorization. The case represents a cultural shift in engineering, including designs that are further customized to purpose, bringing the customer closer to the machine design process, and continuous learning from the existing fleet of cranes. With the digital twin, the traditional assumptions can be challenged by enabling closed-loop design by leveraging data and leading to more optimized designs. Further exploration of usage-based design optimization is described in Section 3.10.

The bearing life estimation case had the most participating organizations: the university and four companies. The crane manufacturer provided the initial motivation for the case, while formalization into a concrete goal was performed as a collaborative effort. The case was defined by looking at the available resources and finding an implementable application from the area of machine design. Most of the practical application development work was done by the university inside the MindSphere/Cloud Foundry IoT platform.

3.8. Case 2: Usage Roughness Indicator

This case features a proof-of-concept usage roughness indicator that shows how smoothly a crane hook is handled. This work was presented at an unrecorded demo session of a project seminar by Valtonen and Ala-Laurinaho. It was also presented as a use case application in a journal article that introduced OSEMA [36]. The application was developed as a collaborative effort between the university and a company member of the project consortium. The university installed a sensor to the crane and connected it to a company cloud. The consortium member performed data processing and visualized the indicator. The components and interfaces of the system are shown in Figure 7.

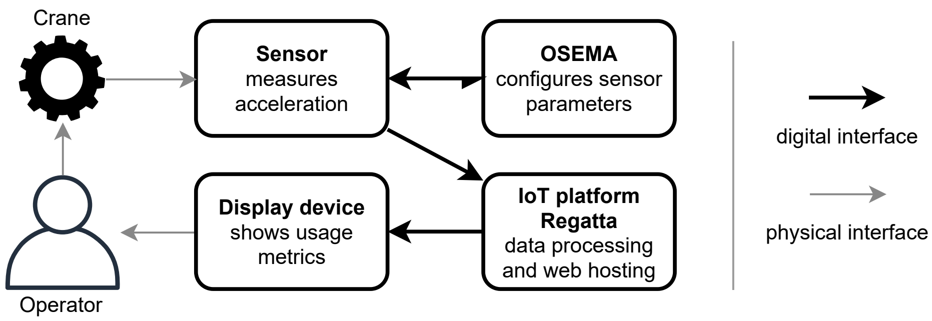


Figure 7. Components of the usage roughness indicator application.

The usage roughness index is calculated from real-time measurement data from a three-axis accelerometer attached to the hook of the crane. The sensor node was configured with an OSEMA instance that was upkept by the university. The sensor sent the acceleration data to an IoT platform of the consortium member. The same consortium member developed a machine learning algorithm to estimate the usage roughness. The algorithm is based on a neural network that was created using Tensorflow. The final application uses a flow-based approach (Figure 8) for data processing.

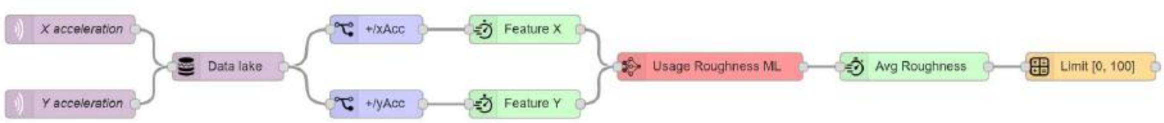


Figure 8. A screenshot sample of the development environment showing the flow nodes of the usage roughness application. The sensor attached to the hook of the crane sends X and Y acceleration data that are processed into an index between 0 and 100.

The machine learning algorithm was developed remotely. First, the crane was operated with different driving styles and the hook acceleration data was collected into the IoT platform. Based on the data, the consortium company implemented a machine learning algorithm. The algorithm was then tested with the crane, and the observations about the performance of the algorithm were sent to the company. Using the feedback, the company improved the algorithm. This iterative loop was repeated several times until the algorithm worked appropriately. Hence, the machine learning model was evaluated to be sufficient through iterative test cycles with linguistic feedback from the crane operator to algorithm developers. A key point of success for the development of the application was an easy-to-use IoT platform, which offered a well-defined interface for receiving measurement data.

3.9. Case 3: AI-Enhanced Brake Condition Monitoring

This case presents a methodology on how brake condition monitoring services can be enhanced with flow-based artificial intelligence (AI). This work was presented in a project seminar by Valtonen and Peltoranta [53]. The case was developed by two consortium members: the crane vendor and an IoT service provider. The case builds on an existing data collection implementation and adds value through cloud-based data processing.

The brake that holds the rope of the crane is a critical component from both functional and safety perspectives. It is also a maintenance-intensive component and hence the crane vendor provides brake monitoring. The pre-existing monitoring services were enhanced with flow-based AI methods to provide deeper insight. As a result, the AI-enhanced twin knows the state of the real counterpart, similarly as people know the state of their health. Based on the health information, the twin provides proactive observations to maintenance personnel. For example, the personnel can be alerted to service the brake earlier than scheduled due to a faster-than-expected wearing speed or to inspect the brake control system if it has triggered an excessive amount of fault signals in a short period of time.

Technical implementation of the brake monitoring consists of "flow nodes" (Figure 9) that perform a certain activity to data and forward the result to the next node. The nodes also act as a visualization of the data processing pipeline and can be grouped into four sections based on their function:

1. Data ingestion creates a time-series data lake by combining the MQTT data channels of multiple flow nodes.
2. Data selection nodes subscribe to topics, propagating notifications of new data to the following processing nodes.
3. Signal processing prepares data for AI with time aggregation, signal aggregation, transform, and other basic math functions.
4. Fuzzy inference (AI) nodes use fuzzy logic to determine crisp indices (e.g., 0–100 %) and linguistic terms (e.g., excellent, good, degrading, worn, critical) for the selected attributes of the analyzed data.

The flow-based AI method provides three types of advantages: (1) Data interpretation: the meaning of data is easier to understand with the linguistic terms. (2) Root-cause identification: e.g., warning in overall brake system health is caused by the wear speed of brake-lining. (3) Predictive maintenance: automatic notices, warnings, and alerts for operators, service personnel, and owners of the crane.

The brake condition monitoring use case leverages the whole physical–digital–physical loop of a digital twin if the human maintenance at the end is accepted as part of the loop: the physical crane generates data that are processed digitally to create insights and alerts that enhance the physical condition of the crane through maintenance actions. In the future, if these algorithms prove reliable enough, critical alerts could trigger restrictions on crane control parameters.

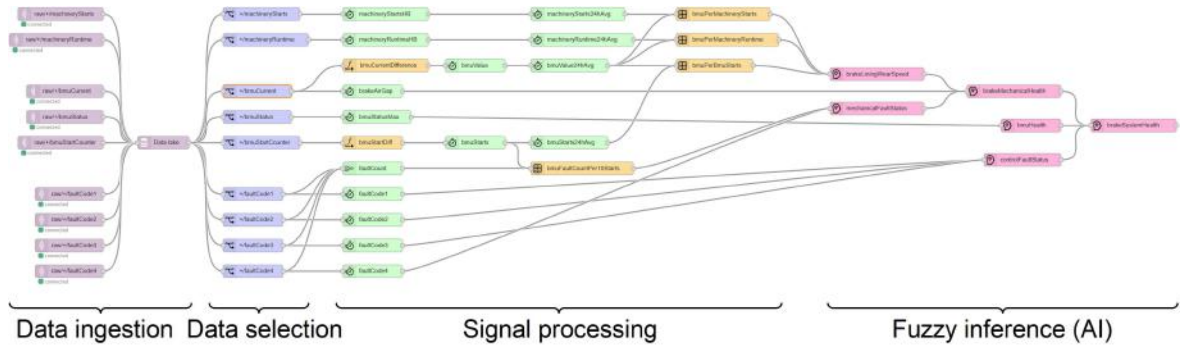


Figure 9. A screenshot sample of the development environment showing the flow nodes of the artificial intelligence (AI)-enhanced brake condition monitoring application. The node categories at the bottom were added afterward. Purple nodes represent data ingestion, blue are data selection, green and yellow ones perform signal processing, and red nodes implement fuzzy inference.

3.10. Case 4: Design Automation

The design automation case presents a proof-of-concept on how real usage data can be used to redesign crane components, specifically a rope sheave and bearing. This case was presented earlier in a project seminar presentation by consortium member Sutela [54] and in a related demo session. The case was implemented by a consortium member with usage data provided by the university.

The application development started with an original design of a combination of a rope sheave and a bearing. Then a design-automation-enabled model was created into a form that can leverage usage data. Usage data collected from Ilmatar was fed to the automated model which created a new design that was optimized according to the usage. The information flow of the case is shown in Figure 10. The results showed a new design for rope sheave-bearing combination whose weight was reduced by 16% and dimensions reduced by 3–10%. The newly selected bearing was 20–30% cheaper.

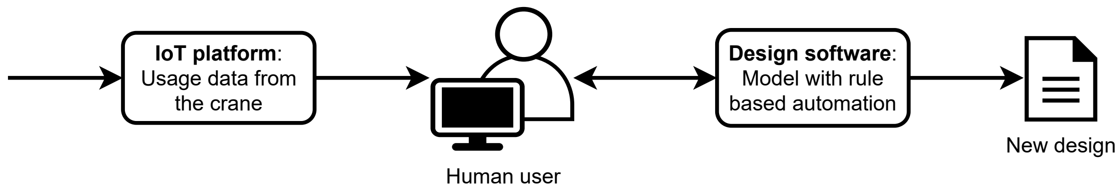


Figure 10. The data flow of the design automation case. The usage data originally comes from the crane as described in Case 1 (Section 3.7).

The case acts as a proof-of-concept of usage-data-driven design automation when implemented with only one crane, but the method can lead to significant benefits when applying to large fleets and more expensive parts. Being able to rely on actual usage data in dimensioning decisions can even enable deviation from standards when they are robust enough to ensure safety.

3.11. Case 5: Web UI with GraphQL

This web user interface (UI) application acts as an alternative user interface to the crane via web technologies, showcasing the potential of the GraphQL wrapper. This case was presented earlier in a master's thesis [38] and conference proceedings [39]. The source code of the web app has been published on GitHub [55].

The GraphQL wrapper is supposed to ease application development for Ilmatar. To demonstrate application development with it, a control application for the crane was developed. In addition to controlling the movement of the crane, the application allows monitoring the internal state of the crane by subscribing to the node values of the OPC UA server. All communication with the crane is performed using the GraphQL interface and, thus, the UI application developer does not need to be familiar with OPC UA.

3.12. Case 6: Mixed Reality Control

The mixed reality (MR) control application shows a set of real-time crane information to the user and allows control of the crane with HoloLens MR glasses. The application was developed as a master's thesis work by Hublikar [56] with Autiosalo as the instructor. The work also includes a prototype of a data linking digital twin to transfer data between the MR glasses and the crane.

HoloLens MR glasses draw 3D images to the user on a transparent screen and allow user interactions via head orientation and hand gestures. The 3D images are stationary in regards to the surrounding world. The developed control application visualizes target destinations for the crane as hologram balls, and when a user taps a ball, the crane hook moves to that location. The glasses also show values from the OPC UA server of the crane, such as x, y, and z position, in a sidebar that is tied to the head movement of the user. Furthermore, a voice control feature was developed and tested in the user test for the application.

The prototype of a data linking digital twin is a program hosted on a separate computer. It reads and writes data from and to the OPC UA server of the crane, and transfers the data via WebSocket to the HoloLens glasses. All the devices (crane, computer, and glasses) are connected to the same local area network via WiFi or Ethernet cable.

The mixed reality control application was developed as a single person project. The development proved laborious with one person implementing both the data link between the devices and the user interface in the glasses. Especially the user interface development environment required a lot of learning and for example, the synchronization of the crane and HoloLens coordinate systems was made only by specifying a fixed initialization location for the app. The source code of the application was not published. Nevertheless, the application proved that a connection between HoloLens and the crane can be established and that the development environment allowed implementing the overall idea. The project also made the pain points of interfaces and coordinate synchronization apparent. Lastly, the user tests indicate that the user-friendliness of the glasses and the overall solution is promising.

3.13. Case 7: High Precision Lifting Controller

High precision lifting controller is a combination of sensors, a minicomputer, and a web application connected to the crane to automatically insert a cylinder (the stator of an electric motor) into a tight hole (the frame of an electric motor). The application was presented by Sjöman et al. [25] with Autiosalo as one of two main developers and continued in a master's level mechatronics project course with Ala-Laurinaho as one of the team members and Autiosalo as an instructor.

The position of the cylinder is measured with sensors that are attached to the frame. Sensors are connected to a minicomputer (Raspberry Pi) which also reads the position of the crane from its OPC UA server. The values are used to control the cylinder to the desired position. The actions are triggered and monitored via a web browser UI.

The OPC UA control interface of the Ilmatar crane was made by the crane vendor from the demand of this project. The development led to the creation of a python library that enabled easy access to the Ilmatar OPC UA server. The library was distributed as an individual file to several other projects until finally published as open-source code on GitHub [57] as part of the launch of Ilmatar OIE more than two years after the initial creation.

The digital twin concept was not considered while developing this case. However, this case initiated the two-way communication interface to the crane, which has enabled a variety of new applications. These experiences are valuable if we want a digital twin to have two-way communication between the crane and its digital twin. Furthermore, the functionality of this case could be integrated into an operator-focused digital twin as an additional feature.

3.14. Case 8: LIDAR-Based Pathfinding

LIDAR-based pathfinding enables automatic scanning of crane environment with a LIDAR sensor and pathfinding to a target with obstacle avoidance. The application was developed at a combined master level mechatronics and automation project course [58] with Autiosalo as the advisor for crane use. The team won an innovation competition with the application [59].

The LIDAR was installed as an additional sensor to the crane and connected via Ethernet to a Raspberry Pi microcomputer which sent the sensor data to a Windows computer via WiFi. The Windows computer performed data processing and sent resulting control commands to the crane with OPC UA python library.

The team was mainly independent in their work and crane support included mainly just handing the documentation of the interface and the crane python library for accessing the OPC UA server. The team noticed an error due to an update and fixed the library. The source code for the application was not published at that time.

This case was not developed for a digital twin and is implemented as a local solution. Nevertheless, this application provides a new local operational feature for the crane. It generates a lot of data about the environment that could be provided to other parties via the digital twin of the crane. The digital twin could also relay control requests to the pathfinding application from external parties, such as a forklift.

3.15. Usage of APIs in the Use Cases

All of the described use cases (Sections 3.7–3.14) are connected to some other building block of the DT as all of them use real data from the crane. The data exchange methods were chosen individually for each use case to fulfill their specified needs. The details of API usage in the cases are shown in Table 1 and the selection processes of the APIs are described in the following paragraphs.

Case 1: Bearing lifetime estimation was developed as an opener case for the project and was planned as a collaborative effort among all project members. The planning phase included an iterative process of finding usable data sources for the crane and finding a topic that serves the goals of the project. The goals were to leverage operational data of the crane from multiple sources and to support machine design activities. Initial ideas included combining information from multiple systems, such as enterprise resource planning and maintenance, but this turned out too ambitious due to practical reasons: access to the systems was limited or they did not have suitable APIs. The case finally leveraged two systems: MindSphere and Teamcenter.

MindSphere was set up to receive usage data from the crane OPC UA interface so it was a natural choice for Case 1. The MindSphere database provided a REST API which was used in a separate web application for two purposes: refining the data and presenting the data. Refining the data required both reading and writing through the REST API and presenting. The REST API authentication (acquiring credentials and learning to use the credentials) required setup, but this proved manageable during the project.

Teamcenter was selected as the target system for the digital twin interface because it was already used by machine designers, had a lot of design data, and was supposed to store product lifecycle information. Teamcenter had a SOAP (Simple Object Access Protocol) API, but it was not used because none of the project personnel had used it before and the time required to learn to use it was considered too high. It was also unclear what the API would actually offer. However, embedding a web browser element to Teamcenter was easy, and this was the method to integrate the MindSphere-based usage data visualization to Teamcenter.

Table 1. Usage of APIs during the development of Ilmatar digital twin.

Case	API Usage	APIs (Highest Open Standard)	API Specification	Used API Libraries	Case Specific Work for API
Case 1	MindSphere gateway fetches data for crane OPC UA server	OPC UA	Ilmatar crane API	-	-
	MindSphere gateway writes data to MindSphere database	HTTPS	Undisclosed	-	Configure new data sources
	Virtual sensor script reads and writes data to MindSphere database	REST	MindSphere	Requests (Python)	Set up authentication and specification
	MindSphere visualization reads data from MindSphere database	REST	MindSphere	Requests (Python)	Set up authentication and specification
	Teamcenter shows an embedded MindSphere visualization	Web embedding	-	-	-
Case 2	OSEMA server updates sensor configuration	HTTP	Custom	Django, socket (Python)	Custom specification
	OSEMA sensor sends data to Regatta IoT platform	MQTT	Regatta	pycom MQTT	Add Regatta MQTT specification to OSEMA
Case 3	Crane data are fed to flow calculation	MQTT	Regatta	-	Additional manual data transfer
Case 4	Data from MindSphere database is transferred to Rulestream	File transfer, csv	MindSphere table structure	-	Manual data transfer
Case 5	Wrapper receives GraphQL request, forwards it to OPC UA server and returns the result	GraphQL, OPC UA	Ilmatar crane API	-	Set the IP address of crane OPC UA server
	Web UI reads and writes data to OPC UA–GraphQL wrapper	GraphQL	Ilmatar crane API	XMLHttpRequest (JavaScript)	-
Case 6	Middleware reads and writes data to crane OPC UA server	OPC UA	Ilmatar crane API	node-opcua	-
	HoloLens application reads and writes data to middleware	WebSocket	Custom	express (node.js)	Custom specification
Case 7	Middleware reads and writes data to crane OPC UA server	OPC UA	Ilmatar crane API	FreeOpcUa (Python)	Created Ilmatar-python-lib
	Web UI reads and writes data to middleware	WebSocket	Custom	WebSocket (JavaScript)	Custom specification
Case 8	Middleware reads and writes data to crane OPC UA server	OPC UA	Ilmatar crane API	Ilmatar-python-lib	Small modifications to Ilmatar-python-lib
	Minicomputer sends data to computer	TCP	Custom	-	Custom specification

Case 2: Usage roughness indicator was developed late in the project after realizing that existing resources and competence could be easily combined into a new kind of application. The application leveraged data from an OSEMA sensor and cloud services of the Regatta IoT platform. Regatta supported MQTT protocol and it was chosen as the method of data transfer as it was recently added to OSEMA and waiting for a use case. The Regatta MQTT specification had to be added to OSEMA, although this proved straightforward as the API was well-documented and the configuration was done by the main OSEMA developer. In addition, Regatta was used by experts in the software, allowing rapid development of the cloud application.

Case 3: AI-enhanced brake condition monitoring was developed using the data from the existing crane and brake monitoring systems. The data for the analysis were extracted from the crane database and streamed into the developed flow implementation (see Figure 9) using MQTT protocol. The flow application routed the data between their nodes using internal, code-level communication routines. The calculated data from the AI-supporting nodes were exported to the IoT platform where it was visualized as time-series graphs.

Case 4: Design automation was built on Rulestream and the usage data of Ilmatar were brought from MindSphere as a spreadsheet file. Hence, this case did not have an actual API. It is included among other APIs to showcase that an API is not always necessary, especially in tasks that are anyway subject to human decision making. The case however used refined data from Case 1 and is an end-user application at the end of an engineering data pipeline that was created with APIs. This acts as a reminder that existing APIs and data can be beneficial beyond the original use case.

Case 5: Web UI with GraphQL was built to showcase the capability of the newly developed OPC UA–GraphQL wrapper. Both were created by the same developer who learned to use both OPC UA and GraphQL during the master's thesis work. Emphasis was put on developing good documentation as the benefits of the wrapper are expected to come from applications built on top of the wrapper.

Case 6: Mixed reality control used the OPC UA python library for the communication between the crane OPC UA server and the middleware. The middleware used WebSocket to communicate with HoloLens MR glasses. Both read and modify operations were used on each interface.

Case 7: High precision lifting controller was the first application to use the crane OPC UA interface for both reading and controlling. The python OPC UA library was developed during this work and it was built on top of an open-source library "FreeOpcUa". The server application used WebSocket to communicate with the client browser.

Case 8: LIDAR-based pathfinding used the OPC UA python library for communication. Aside from this, it uses non-standard or local communication methods leveraging C# and TCP.

3.16. Publishing the Results from the Crane Environment

The publicity combined with concrete actions and coordination makes the Ilmatar environment a special industrial research platform. The crane is a tangible device for which it is easy to innovate new applications, and coupled with coordination from university and industrial partners, Ilmatar sees an extraordinarily large amount of development activities. Hence, it generates an extraordinarily large amount of (qualitative) product development data. In distinction to usual corporate research equipment, a significant portion of Ilmatar data is public. Most of the published data are in academic publications, such as conference papers [24,39], journal articles [25,36], and master's theses [38,56]. Some of these works were purely descriptive with no actual application, some featured an application that was not published, and for some, the software was published as free open source in addition to the academic publication.

Publishing pieces of software as open-source code creates extra unrewarded work in most cases. However, there has been time savings from publishing one piece of software. The python library developed for the crane during high precision lifting development (Section 3.13) was useful also for future projects, but distributing and maintaining an updated version of that library required effort. Publishing the library as open source [57] removed the need for separately distributing the software and streamlined user contributions to the library. As a downside from the open sourcing, it is not clear now who uses the software which would help keep track of the demand for the library and collect user feedback. (This is one of the reasons why the Ilmatar OIE resources were put behind registration.) The other open-source projects made in the crane environment have been published without user demand and have not yet attracted contributions outside the original creators. They are more complex than a simple library and they are not as obviously required as the crane library, which directly takes away work from those who want to connect to the crane OPC UA interface.

3.17. Scalability of Digital Twins in Company Operations

Here, we present observations on how digital twins can be taken into a normal part of company operations. The contents of this subsection are based on consortium member Lehto's presentation at a project seminar [60] and they are included as base material for further discussion.

Each of the presented cases shows potential benefit for different stakeholders of industrial cranes, but a question about payback time remains as creating a digital twin requires extra effort instead of a situation where a digital twin serves its users right from the beginning. Once digital twins become more integrated with product development, creating virtual products will likely be more cost-efficient and contain more synergies than today.

There cannot be a project for every digital twin a company produces. Digital twins are scalable only through a plug-and-play implementation that is supported by IT systems across the company. Company-wide adoption can be achieved through a strategy for digital twins and IoT data. This strategy determines what data are important, to whom, and why.

With a strategy that supports digital twin creation, a company can determine whether benefits outweigh costs. The costs come from multiple activities during the DT lifecycle, such as creation, sensoring, DT structure upkeep, updating the service and modernization actions, a data stream from a physical environment, analytics, computational power, and distribution of information reports. Depending on the use case, DTs may include additional features that induce costs outside these categories, and some costs can be minimized through automation.

Usages for widely implemented fleets of digital twins include not only the previously mentioned new product development and maintenance services but also for example sales who can use the existing fleet data to show what kind of cranes fit what kind of customer applications. As data become more easily available across the company, we see it probable that more use cases emerge, and these benefits come "free" on top of the already alleviated costs. This can create a positive spiral, leading to a situation where digital twins are expected to become a normal way of handling any information across the whole company.

The company-wide culture change from using familiar local data to using new cloud-based DT data is undoubtedly a major challenge, but it is crucial for achieving all the potential benefits. Digital twins and their interfaces need to be integrated into and serve the everyday company processes.

4. Discussion

The discussion section presents the insights, lessons learned, and recommendations from the development experiences described in Section 3.

4.1. Integrated Digital Twin

The concept of an integrated digital twin developed gradually from various phenomena observed during the creation of the Ilmatar digital twin. The basic conceptual ideas were presented in an earlier publication [16] whereas the current paper shows a practical DT implementation and formalizes the concept of integrated digital twin according to these experiences.

By an integrated digital twin, we mean that all components of a digital twin are available from a single web location as seamlessly as possible. Instead of forcing users to fetch information from several separate systems, an integrated twin brings the information to users based on their needs. An integrated digital twin can be used by software agents and human users as depicted in Figure 11. Therefore, integration means either machine-readable APIs or human-perceivable views to information. The human view is provided by a UI application that is connected to the machine-readable interfaces, highlighting the fact that DTs operate in cyberspace which is not native to humans.

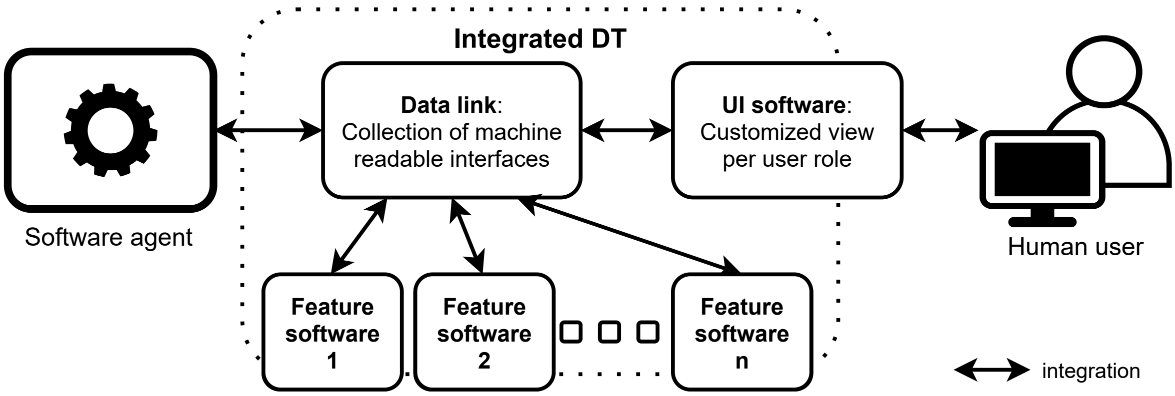


Figure 11. Services and users of an integrated digital twin. The data link is not software at its core but is made available via software. "Feature software" blocks represent the features of feature-based digital twin framework (FDTF). The figure does not depict the connection to the real-world entity, which is fulfilled by one or many "Coupling" feature software blocks.

The borders of the integrated digital twin are depicted with dotted lines in Figure 11, making it unclear if the features are a part of the DT or not. This is a conscious choice because during the development it proved impossible to judge if a selected software is part of the digital twin. Any attempts for hard lines seemed arbitrary and were dismissed. Instead, an integrated digital twin is defined by the connections between software blocks:

> An integrated digital twin is a collection of digital services available via a single digital location, linked to a single real-world entity.

This services focused definition avoids the ambiguity related to determining if a software is part of DT. The services are provided by the software blocks which are part of the digital twin service supply chain. In the cases of Ilmatar, each block is crucial in providing the end result and is therefore an obligatory part of the service supply chain. However, each case was built as an independent whole and some have manual activities as part of the data supply chain. Work needs to be done before the individual DT applications are combined into an integrated digital twin. Although in the long term, there should be rather a natural push towards integrated DTs, as a digital twin should be a service for both crane users and application developers. The need for a platform for building integrated digital twins is evident.

Reaching the level of an integrated digital twin already seems to be a general unsaid goal for any digital twin development project. To demonstrate this phenomenon, we describe two cases presented in the Aalto environment. Another university research project developed a digital twin for a rotor system [61]. This reached a higher level of integration into one digital twin with a web-based 3D view of the rotor which combined the other features, including sensor data from two sources and neural network-based virtual sensors. A simple broker server software was developed to combine data sources with customized data adapters where necessary. The twin was built almost entirely at the university, minimizing the need for cross-organization collaboration. The rotor system serves as a good example that with good coordination it is possible to build integrated digital twins, in this case for research equipment built with industrial-grade components.

Another example of an integrated digital twin was shown by a mining technology and plant supplier company Outotec in a demo session of the project seminar. They used Aveva software (AVEVA Group plc, Cambridge, UK, https://sw.aveva.com/digital-twin) to create a "plant engineering digital twin" that brought engineering information of a processing facility available via one 3D model. The twin integrated information from multiple types of documents, such as layouts, lists, and diagrams, into one view, and is a good example of an integrated digital twin used in the industry.

Even with the encouraging examples from building integrated digital twins in single organizations, current tooling seems inappropriate. Creating properly integrated cross-organizational digital twins requires so much additional labor on top of the actual application development that the job simply does not get done with a decent amount of resources. Hence, a new coordinated approach for building integrated digital twins is needed.

We propose creating a digital twin platform that is independent of any feature software or providers. We also propose two design principles for the platform: openness and user-centered design. Openness is further divided into two components: open source software and open standards. Openness because digital twins should be located on such a low level in the technology stack that the required network effects will only happen with an open solution. The World Wide Web and containerization are good examples of Internet-based technologies that have created network effects via an open approach and we see that digital twins should be located on a similar level of the Internet stack. User-centered design is required to ensure adoption of the DT platform, including good usability for both digital twin end users and developers as users of the platform. The goal is that a digital twin naturally attracts content because it is the handiest place for it, and adding a new feature to a digital twin becomes as easy as downloading an app to a smartphone.

4.2. Hypotheses Review

We now review the eight hypotheses shown in Section 1.1. The review is based on the experiences of one environment during a relatively short time, and should therefore be taken as indicative rather than conclusive.

> Hypothesis 1: Digital twin transforms data from a physical product to useful knowledge. Digital twin offers data and knowledge to all stakeholders across the product lifecycle.

The bearing lifetime estimation and brake condition monitoring case focus strongly on turning data to knowledge through multi-phase data processing. The usage roughness indicator also turns raw data into more sophisticated information about how the crane is being handled. The design automation case uses existing knowledge of the application and turns data into an actionable solution. The web UI shows data rather than transforming it into knowledge. Hence, turning data into knowledge seems to be crucial in some cases, but not in all. This emphasizes that digital twins are defined by their use case, and different features of DTs have different purposes.

The Ilmatar DT as a whole provides the data or knowledge to designers, maintainers, and users of the crane, although each case focuses on just one stakeholder. Currently, these cases are fragmented, and serving all stakeholders from one digital twin demands a lot of integration. Both of the claims of hypothesis 1 can be fulfilled if the digital twin is a combination of multiple cases and they are often perceived as goals for twins, but they are not general requirements for digital twins.

> Hypothesis 2: Digital twin integrates digital models and data from different sources and providers and offers a customized view for each stakeholder.

The models and data used for the components of the design and maintenance focused DT cases come mainly from the manufacturer and the crane itself. The operator focused applications heavily rely on crane data, but also leverage external data, such as the additional sensors of the high precision lifting case and the environment data of the LIDAR-based pathfinding case. Hence, it seems that design and maintenance can rely on data from a limited amount of sources, whereas operation focused applications also crave other data sources. Although, it may also be that we did not find the right supplementary data for design and maintenance purposes. Integrating the work to a single digital twin proved to be more difficult than expected, but is still seen as a prominent development direction.

The existence of customized views is initially verified, as the components of Ilmatar digital twin offer views for machine design and maintenance as well as operators. In addition, sales have been identified as another relevant stakeholder. Each of these groups needs a customized view to achieve their goals.

Hypothesis 3: Digital twin enables networking and business.

Networking has several points of validation across the use cases, although they come with two preconditions: functional interfaces and a commonly beneficial use case. The digital twin concept itself as well as the buzzwordiness of the term acted as business networking enablers. Actual business creation is more difficult to validate in the context of this externally funded research project as most of these were made as proof-of-concept prototypes instead of business-critical applications.

The usage roughness case had clear API-based boundaries of responsibility between the two organizations, which enabled efficient information exchange. The clear boundaries can be thought of as a networking enabler thanks to the efficiency of communication they offer, and we find it probable that if taken to a business context, the API-based responsibility boundaries will enable easier business creation.

The bearing use case combined most parties to one case: one company provided a use case, the university acted as an integrator and application builder, a second company provided analysis algorithms and the end application was integrated to the PLM software upkept by a third company. The application platform was provided and the data stream setup by a fourth company. Making an integrated digital twin application naturally combined parties into a digital supply chain network after a common goal had been defined.

On the conceptual side, the digital twin provided a common goal for the project consortium. Each participant had their role in that vision and even though they weren't all directly linked to each other, aiming to create one digital twin with several features brought the network together. We are expecting that a more integrated digital twin will support even stronger networking when the results are combined to one interface.

Buzzwordiness around the digital twin concept is another aspect that seems to create a lot of networking, for example in the form of the number of people participants in events. Each of the two seminars organized for the project attracted more than a hundred participants mainly from industry, an exceptional amount for such a small nationally-funded research project. The digital twin term gathers together people from various disciplines to pursue a common goal: creating digital twins.

Hypothesis 4: Machine design dimensioning and product development processes can be redefined with the true usage and maintenance data provided by a digital twin.

One of the project partners fed usage data to a design automation model to create usage-based dimensioning of the rope sheave. This approach works only in hindsight but can be developed further to achieve redefinition of existing processes. If the whole crane is designed with rule-based design automation and the usage profile of the customer can be estimated accurately (for example based on data from similar customers), cranes can be designed and manufactured to fit the expected use case more accurately, allowing lower overall expenses. In addition, if the usage of machine parts can be measured, such as for bearings and brakes of Ilmatar as described earlier in cases 1 and 3, lifetime estimations of those parts can be developed to become more accurate and parts can be changed when their usage reaches their designed lifetime, not by trying to estimate usage or looking for signs of failure. True maintenance data was not included in the Ilmatar digital twin so far, but it will be needed to make accurate estimations.

> Hypothesis 5: The overhead crane located at university premises acts as an excellent development platform, offering industrially relevant applications, such as plugging digital twin as part of product configuration, design, and life-cycle management.

The experiences for this hypothesis are mixed, and depend on the viewpoint. From the university point of view, the applications are very industrially relevant, but from the industry perspective, the cases were rather academic proof-of-concept tests and the results could not yet be implemented in company operations. The uniqueness of the crane also created some obstacles in development, as the usage profile of Ilmatar differs from most cranes. Nevertheless, the crane has proved to be a unique development platform, enabling research collaboration in ways not possible earlier. When the pros and cons of this type of environment are known, research and development activities can be planned to take place on topics that are supported by the environment. For example, data access to the crane was comprehensive and it was easy to share it and generate small amounts of targeted usage data, but the long-term usage data does not match those of high-usage industrial cranes. The public nature of university and integration to teaching should also be considered when planning development activities. The general outlook on the environment has been positive and anticipatory for future developments.

> Hypothesis 6: Acting as an interface for all Industrial Internet data is one of the most important functions of a digital twin, enabling the efficient use of a vast amount of data.

The need for this kind of integrated digital twin grew during the digital twin development, but it proved to be such a complex task that no concrete evidence was acquired to support that this is an important task especially for a digital twin. Easy-to-use interfaces in general proved to be a very important enabler, as they seem to speed up application development significantly. The conceptual frameworks developed during the project support this hypothesis.

> Hypothesis 7: "Using existing APIs enables fast prototyping, and bringing 'developer culture' from the 'software world' to the 'physical world' enables faster prototyping/product development cycles" [25].

This hypothesis is supported by all cases and two cases proved especially supportive for faster prototyping. The development of the Web UI with GraphQL was trivially fast thanks to the OPC UA–GraphQL wrapper. The development of the usage roughness indicator was fast as the used IoT platform provided a well-defined interface to which the OSEMA sensors were easy to connect.

All but one of the operator-facing applications were built on top of the pre-existing OPC UA API, and the usage roughness indicator was built on an existing IoT platform and the pre-built OSEMA sensor platform. From the design and maintenance focused cases, bearing lifetime estimation and design automation also leveraged OPC UA via the MindSphere IoT platform and the AI-enhanced brake condition monitoring used an existing interface and IoT platform.

However, most of the applications are focused on user interfaces with no changes to physical products. The design automation created new designs for crane parts, but the results were not used for actual physical prototyping. We also had the API of only one crane, while making proper conclusions for physical development would have needed APIs from several cranes. It also seems that many current APIs are too laborious to be used efficiently in iterative physical product development.

> Hypothesis 8: Digital twin can be built without selecting a central visualization and simulation model.

We built the digital twin without leaning on one visualization and simulation model and it came out fragmented in multiple separate cases. We see that this approach allowed a wide exploration of different features for digital twins, but we ended up with an intangible result: we have several separate instead of one integrated digital twin of the crane. However, we still see this as a direction worth exploring, and to overcome the difficulties, we are building a "data link" tool to tie the pieces together more concretely without model-based visualization. Hence, the hypothesis seems valid, but extra care should be put into combining the different pieces together if a visualization model is not used.

In conclusion, most of the hypotheses were validated at least partially. Some of the hypotheses proved too ambitious to be fulfilled in one project, but are seen as prominent development direction.

4.3. API-Based Business Network Framework

Observing the collaboration of different stakeholders of the project led to the discovery of a framework that states that business networks should be designed in parallel with the corresponding technical application network. The approach is to leverage technical interfaces (APIs) as the basis for organizational boundaries in a business network. More specifically, when organizations collaborate, they have both a technical API and a business relationship, and the structure of these relationships is built as identical as depicted in Figure 12.

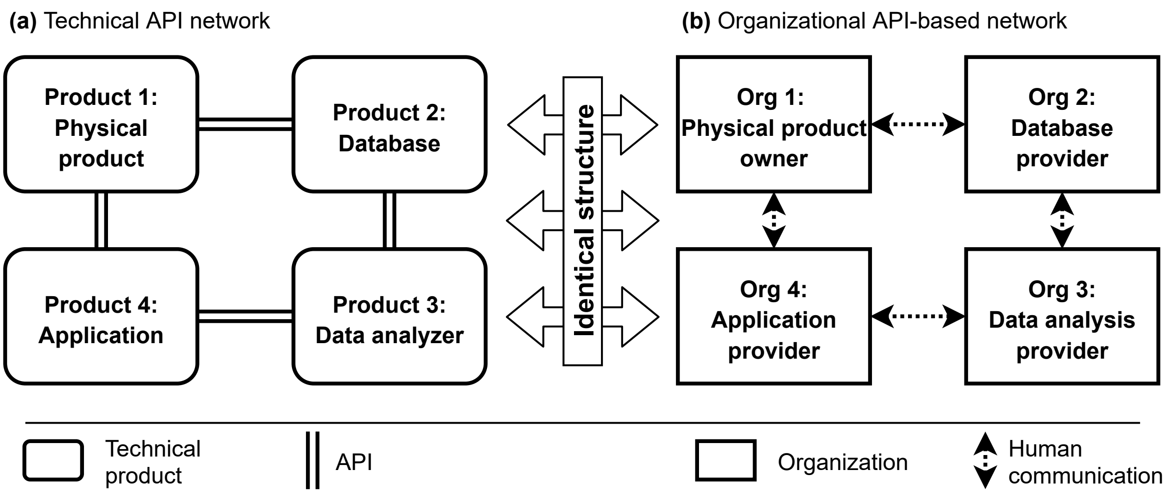


Figure 12. Example illustration of API-based business network framework. The data from Product 1 owned by Organization 1 go through a loop via digital products before providing an application for the physical product. The structure can be of any form and does not need to be a loop.

Our hypothetical theory is that the structure of APIs between organizations should be used to define the structure of teams and their responsibilities in digital twin development. This makes sense because when data are moved to a service managed by another team, the responsibility of working on that data changes. Each collaborator works on their own platform and on the data they receive from others (or generate their own data).

There are multiple projected benefits for this kind of network design, mainly in enhancing communication. A relationship that is built around a technical API is simple and traceable; the API either works or not and it either gives the specified data or not. The data specs are specified in the API documentation. It is immediately visible if the API stops working or gives bad data. Hence, it is easy to identify who should fix the issue. The API-based network architecture also contributes towards the general data monetization [62] ambition. When a business relation and an API are parallel, it should be natural to define monetary value for the data that is transferred through the API.

The distribution of ownership for the different parts of the application is an intended design feature of this style. This is becoming necessary as applications are required to be so complex that it is not sensible for the main application owner to master all the technical details of the application. Instead, the application owner becomes an API architect who does not need to see what happens behind each API; it is enough that the input and output of each block work as agreed. These factors lead to the basic characteristics of API-based business networks:

- The ownership and boundaries of services are obvious.
- APIs are the primary design tool of an application architect.
- The inner workings of services are abstracted.
- All service providers must work with APIs.
- The competencies of more teams can be included in the application.
- Each service can be scaled individually based on demand.
- Changing service provider becomes straightforward.

The API-based business network architecture style highly resembles the microservice organizational style commonly used in web application development. However, industrial DT services have physical devices as parts of the application, which brings in more diverse computing environments, more complex supply chains, and long support periods. The diversity of computing environments appears in several forms, such as a high number of operating systems, limited computing power, limited connectivity, and additional security demands. The complexity of supply chains for physical products is high as a large number of parts are bought from subcontractors and the goods need to be shipped physically. Long support periods are required for physical products whereas IT products can be even completely changed. Therefore, while this API-based organizational style has been used in software-only, special attention is required to make it work also for industrial DT products.

The API-based networking style can be demonstrated by two examples from the project. The usage roughness case (Section 3.8) provided inspiration and a well-working example of the network style. The two parties relied their cooperation on data exchange through a single API; one party provided the data and the second analyzed it and provided the visualization application. If modifications were needed, requests were made via email, but the data were still exchanged via the API. Additionally, the whole Ilmatar DT can also be visualized according to API-based networking style as shown in Figure 13. The identification of organizations was made after the completion of the project, but the components still have clearly defined owners. Based on this project it seems that ending up with the API-based networks style structure is a natural way of organizing a multi-component industrial DT. Therefore, it seems logical to use this framework as a design tool for DTs.

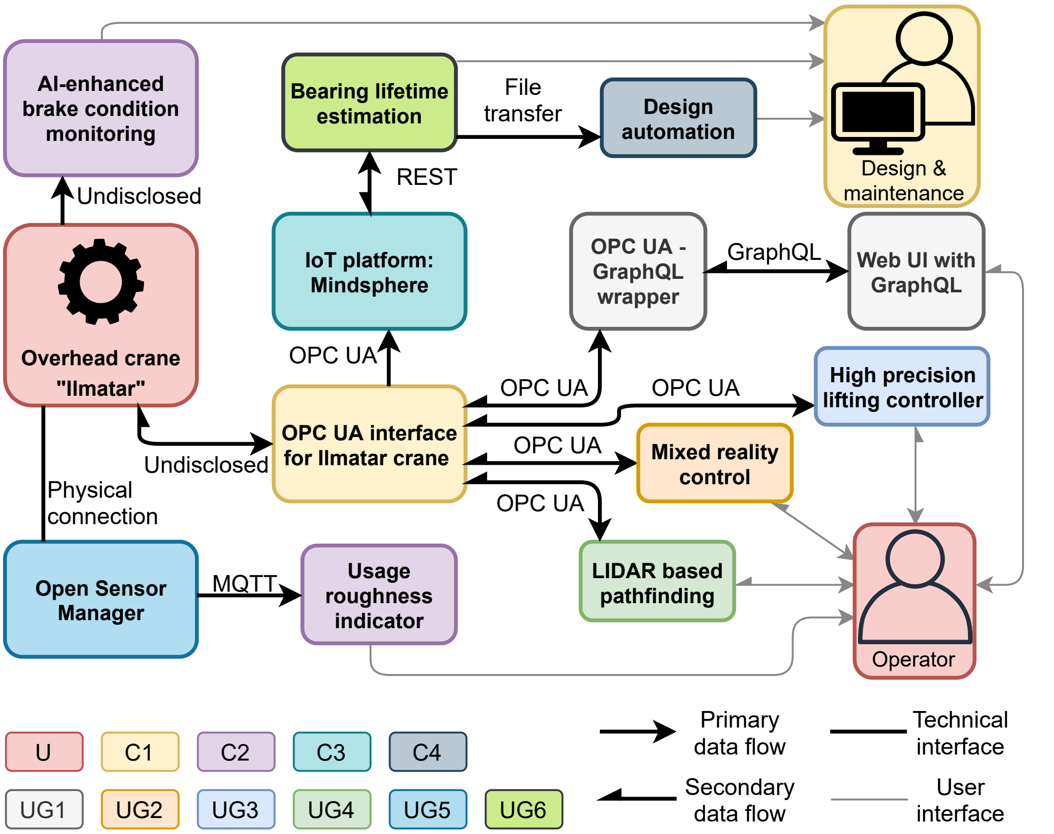


Figure 13. A case visualization of the API-based business network framework: The components of the Ilmatar DT are colored by the organization in charge of their operation. U: university in general, C1-C4: companies, UG1-UG6: university groups.

The API-based business network architecture framework is currently a hypothetical theory and requires further experimentation before benefits in operational industrial environments can be verified. Practical implementation requires all participating organizations to be invested in APIs, more than the current industry standard. However, it seems inevitable that more and more companies start using APIs as part of normal operations and in this kind of future, it is only natural to end up with API-based business networks. At the current form, the framework can be used as a communication tool to present and plan multi-organization DTs efficiently.

In a related study, Barricelli et al. [10] brought up the complexity of collaborative DT design projects, emphasizing the communication and skill gaps between participants from different domains. Barricelli et al. encourage using a sociotechnical design approach to ease communication gaps and to enable development also for experts outside the IT domain. This approach seems to support the need for the API-based business network framework and requires using APIs that are user friendly also for people without an IT background.

4.4. Lessons Learned

During the study, we recognized the following six themes that should be given special attention when developing digital twins that have more than one component.

APIs. Various types of exchanging data between systems proved to be an area with a lot of space for rapid development. APIs have been a basic tool in IT systems development for decades, but the DT world is only on the verge of recognizing them. There is a multitude of different API types in various dimensions: they can be local or remote, private or public, just a library, standardized or unstandardized, or between these categories. Documentation can often be understood only by IT experts. Hence, knowing everything about the relevant APIs seems impossible. Nevertheless, we made a basic flowchart for selecting APIs during DT application development based on the experiences of the project, shown in Figure 14.

Key takeaways from the project are that APIs should be used more in DT development, any properly implemented API services become a part of the company infrastructure that can be used in later projects, and APIs need to be easy to use for as many people as possible to be used efficiently. Further notes on the usability of APIs are given in Section 4.5.

Standards. During the project it became apparent that a standard for describing digital twins on metadata level would be needed to make extendable digital twins. The digital twin of the crane consists of several parts built for different purposes with different tools, but there was no method on how to state in a machine-readable format that all these parts belong to the digital twin of our Ilmatar crane. Standards for metadata exchange are now being developed [63].

Tools. While there are tools for implementing specific components of digital twins, there were no tools for combining these. In addition, only some of the component tools can be launched as services with 24/7 availability, which is a practical requirement if the component is leveraged by other teams. Furthermore, launching even simple web services proved to be unnecessarily difficult during the study. Therefore, we have two recommendations for digital twin tool development: digital twin builders that combine multiple components and making it easier for non-coders to deploy the components as 24/7 services. A recently published open-source tool "Digital Twins Explorer" by Microsoft [64] is a good opening in the right direction.

Open source. The majority of the cases of the project were built on top of open-source solutions: all the cases built by "University groups" in Figure 13 leveraged open source and for example, the MindSphere IoT platform is based on open-source software. Open solutions can be tested instantly and therefore they seem to be especially suitable for innovation. Internet and software companies have found ways to leverage this innovation potential by both using and offering open solutions, whereas industrial companies are still reserved towards openness. Digital twins are mostly software and leverage the Internet, so any company building them should get familiar with open solutions to stay competitive. It is also good to acknowledge that open source is a complex field: for example, there are a plethora of open-source licenses, multiple business model styles, and specific dynamics for community engagement.

Skills. The digital twin applications of the project were developed mainly by mechanical engineers, although some of them were at least moderately experienced in programming, which proved to be an essential skill in several DT components. Even though creating many advanced features is certainly possible, they may not be implementable in a limited time frame. This can be especially deceiving when planning to use new tools and the expectations for those tools don't match reality. On the other hand, the right tools can also remove the need for experts in certain areas. Each digital twin development project should critically evaluate if they have the necessary combination of skills and tools available for developing the type of digital twin needed. Skills have been recognized as a crucial factor also by other researchers [4,14].

Goal. Developing multi-component DTs is currently an uncharted territory with no standard implementation guides, which means that project participants need to both apply their knowledge in new ways and learn new skills. This lack of best practices and example solutions makes it practically impossible to plan the detailed outcome of the project in advance. Hence, it is important to define, communicate, and update the goal of the project constantly as new skills are acquired. The two methods for goal management during the project were to define a purposeful use case and to practice cross-organizational leadership, which are further described in Section 4.5.

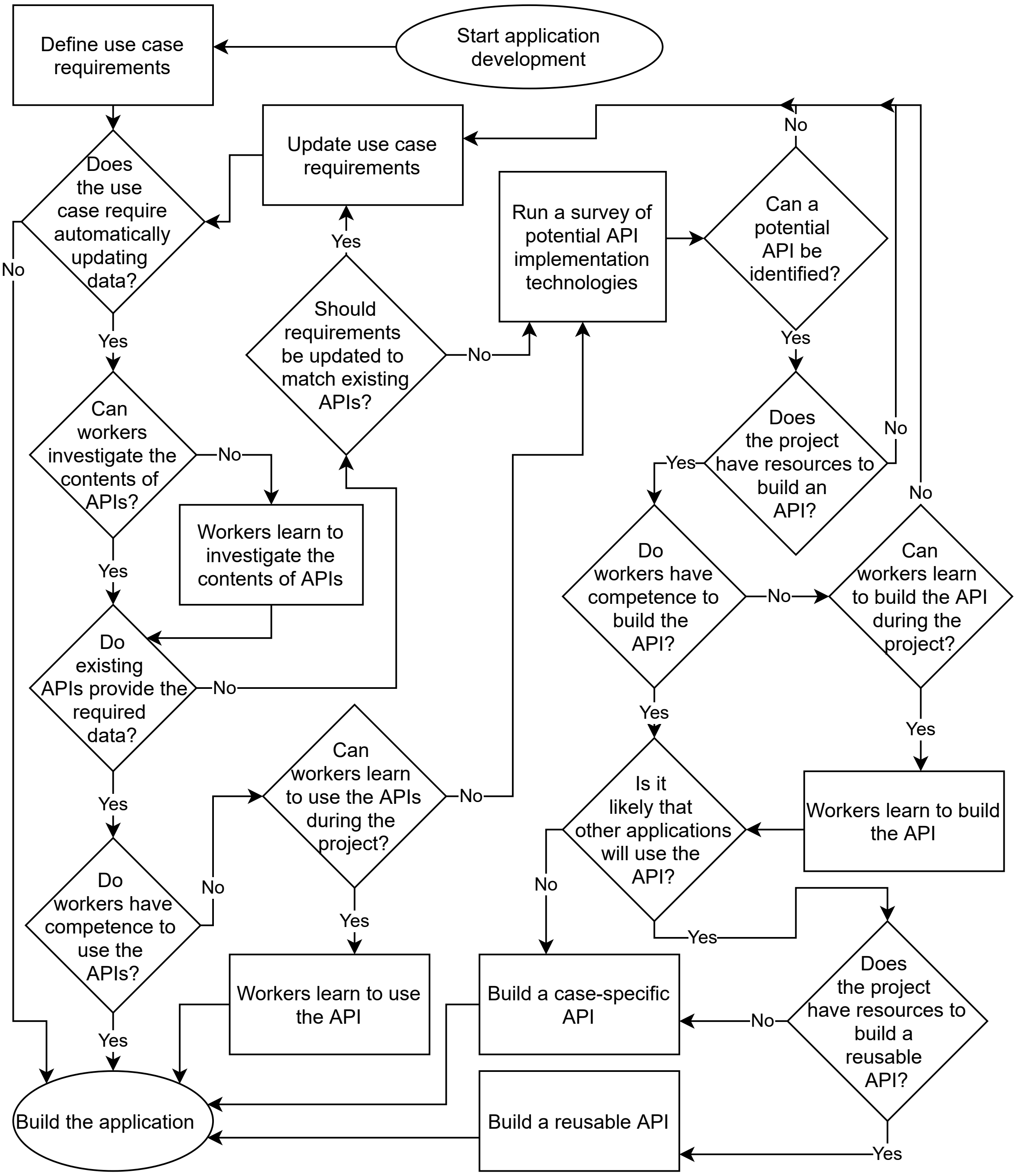


Figure 14. Flowchart for acquiring API resources for a digital twin (DT) application development project as observed during the example project. The diamonds are fairly quick checks, whereas the rectangles represent various amounts of work. (e.g., building a reusable API requires more work than building a case-specific API.) Hence, rectangles should be avoided to get started with actual application development as fast as possible. Notes: The word "workers" refers to all the project workers collectively, i.e., it is enough that one of the workers has a specific skill. As an alternative for workers learning new skills, the project can also use outside support to perform the consequent tasks, although this may lead to a prolonged dependence of the support and possibly including the support in updating the use case requirements. Defining and updating the use case requirements is included in the chart as they proved to be inseparably dependent on the availability of suitable APIs.

4.5. Managerial Implications

The most important takeaway from the development of digital twin applications for Ilmatar crane was the significance of easy-to-use APIs. They seem currently undervalued and leveraging them properly would offer significant efficiency boosters. Crucial factors for achieving integrated digital twins include also a purposeful use case and cross-organizational leadership. In the long run, the role of open standards and open-source software need to be acknowledged and developed intentionally.

Easy-to-use APIs are important because they are a requirement both for building integrated digital twins and for browsing digital twin data. Ease of use is of course a subjective matter, varying per person, and it is important specifically as such. The APIs need to be usable for those who need the data in their work; it does not help if they are usable for the most experienced developer in the company. Unfortunately, current methods to view the data from APIs are often cumbersome. For example, the user interface of the popular API browser "Postman" (PostMan, Inc., San Francisco, CA, USA, https://www.postman.com/) relies heavily on typing text instead of offering clickable buttons. This text-based UI may be preferred by developers thanks to its versatility, but it is impractical for someone with little programming experience who just wants to browse the data behind the API. Contrastingly, the OPC UA client "UaExpert" (Unified Automation Gmbh, Kalchreuth, Germany, https://www.unified-automation.com/products/development-tools/uaexpert.html) has a graphical interface that allows browsing by clicking through a tree-like structure. MindSphere offers a graphical web user interface for browsing the historical data. We recommend decision-makers to demand this kind of usability from all API providers (both internal and external). To test usability, you can simply ask yourself: Can you browse the data behind the API? If not, probably neither can the majority of your employees, which makes developers a constant bottleneck for data access throughout the organization. APIs offer a whole new view to operations and enable innovations, but only if they are accessible. We recommend treating APIs as investments.

A purposeful use case is a starting point for the actual digital twin development. The digital twin development for the Ilmatar crane started out by defining a common use case, which led to the formation of the bearing lifetime estimation case that had roles for all but one of the project partners. Other cases were defined among fewer participants and led to isolated twin applications. Combining these all to one twin proved both technically difficult and lacked purpose. It should also be noted that use cases were defined by what is possible with the currently available APIs, further highlighting their importance.

Cross-organizational leadership is required both when planning the use case and implementing digital twins applications that cross organizational boundaries. The planning stage requires insight into the competencies of all participating organizations. During implementation, cross-organizational leadership helps to stay focused or to make the decision to change plans. Cross-organizational leadership is important for integrated digital twins because they combine competencies and products of organizations in ways that differ from their old practices. The importance of leadership can be reduced by having strong interfaces and a clear purpose, or even by following the API-based business network framework (Section 4.3) so that everyone knows their role.

Open standards form the basis for the World Wide Web and it seems that the network of digital twins will only materialize with a similar approach. It is important to distinguish freely accessible open standards and paywalled traditional standards as there seems to be a conflict between the proponents of these styles. The traditional standards may be overlooked by developers if they cannot be accessed easily, and some industries may think the Internet standards are not standards at all. Digital twins need them both in their mission to merge the physical world with cyberspace.

Open-source software is often overlooked by industries due to various reasons, often unnecessarily. While there are cases for both open and closed source software, the unique benefits of open source software, such as community creation and developer friendliness, should be taken into account when developing the company software portfolio. Many companies also use open source more than they think. For example, a vast majority of Internet servers are run on Linux, so it is not a question of "Does your company use open source software?" but rather "How much open source software does your company use?" When you combine the facts that software companies are already heavily relying on open source and that machines are becoming increasingly software-based, it seems inevitable that manufacturing companies will start using more open-source software. It is time for manufacturing companies to start preparing an open-source software strategy.

4.6. Limitations

The method of data collection used in this study (Participatory Action Research) is known to pose a risk of researcher bias. In an attempt to alleviate this risk and to pursue objectivity, we presented also matters that did not go well during the studied project. In fact, the main conclusions are based on the encountered difficulties. Nevertheless, the fact that the authors participated in the project may render them blind to some aspects of the project that could be seen by outsiders and this should be acknowledged when reading the study. However, this data collection method allowed authors to acquire deeper data than is possible through outsider observation.

The method of theory formation used in this study (Grounded Theory) concentrates on new theory generation rather than theory evaluation. The conclusions of this study are supported by a limited amount of data, i.e., one industry–university project, and should therefore be later be evaluated against more development experiences to see if the observations were specific to the research project, or applicable to the development of integrated digital twins in general.

5. Conclusions

This study presented, analyzed, and gave recommendations based on an industry-university project that developed a multi-component digital twin for an industrial overhead crane. The twin is built with two tools (OSEMA and OPC UA–GraphQL wrapper) and three frameworks (FDTF, DT core, and DT-PLM) developed during the project and consists of eight separate application cases built for the designers, maintainers, and operators of the crane. One use case was developed as an integration of multiple systems and stakeholders, but the rest of the applications were not merged into one coherent digital twin. It became clear that building integrated digital twins demands a lot of coordination work that may not be worth the trouble with current tools. Hence, the current lack or unsuitability of tools is seen as a major barrier to the development of integrated digital twins.

The cases indicate that user-friendly APIs speed up application development and are even a prerequisite for the innovation of applications. However, leveraging the current APIs efficiently requires new skills from the workforce: first, an overall understanding of what can be achieved with APIs should be attained by every employee, second, the technical know-how to use them as tools for those who can benefit from API data in their daily work, and third, the technical skills to provide APIs as service to other employees. Currently leveraging API data demands too much work, so we recommend investing in user-friendly interfaces and treating them as valuable digital infrastructure.

We formalized the concept of integrated digital twins and reviewed a series of eight digital twin related hypotheses with mostly supporting results. We also present experiences from making the research environment a public innovation platform. We discovered a novel API-based business and innovation network architecture style as a way to structure digital twin data supply networks.

In conclusion, our development experiences indicate that we should continue striving towards integrated digital twins while more user-friendly tooling is required before this can be achieved.

Author Contributions: Conceptualization, J.A., M.V., V.P., and K.T.; methodology, J.A., M.V., and V.P.; software, J.M., R.A.-L., M.V., and J.A.; resources, J.A., R.A.-L., and M.V.; writing—original draft preparation, J.A. and R.A.-L.; writing—review and editing, J.A., R.A.-L., K.T., and M.V.; visualization, J.A., M.V., and J.M.; supervision, K.T.; project administration, J.A., V.P., and K.T.; funding acquisition, K.T. and J.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Business Finland grant number 8205/31/2017 "DigiTwin" and grant number 3508/31/2019 "MACHINAIDE".

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The processed data presented in this study are included in the article and cited works. Restrictions apply to the raw data of the development process due to confidentiality. Additional information is available from the corresponding author upon reasonable request.

Acknowledgments: The authors would like to thank all DigiTwin consortium members and those who presented or participated in project seminars. J.A. would like to thank KAUTE Foundation and Walter Ahlström Foundation. R.A.-L. would like to thank Tekniikan edistämissäätiö.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	artificial intelligence
API	Application Programming Interface
DT	digital twin
DT-PLM	digital twin-based product lifecycle management
FDTF	feature-based digital twin framework
IoT	Internet of Things
IT	Information Technology
LIDAR	Light Detection and Ranging
MQTT	Message Queuing Telemetry Transport
MR	mixed reality
OIE	Open Innovation Environment
OPC UA	Open Platform Communications Unified Architecture
OSEMA	Open Sensor Manager
PAR	Participatory Action Research
PLC	Programmable Logic Controller
PLM	product lifecycle management
REST	Representational State Transfer
SOAP	Simple Object Access Protocol
UI	user interface

References

1. IEEE Computer Society. IEEE Computer Society's Top 12 Technology Trends for 2020. Available online: https://www.computer.org/press-room/2019-news/ieee-computer-societys-top-12-technology-trends-for-2020 (accessed on 15 July 2020) .
2. Grieves, M.; Vickers, J. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In *Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches*; Kahlen, F.J., Flumerfelt, S., Alves, A., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 85–113. doi:10.1007/978-3-319-38756-7_4.
3. Shafto, M.; Conroy, M.; Doyle, R.; Glaessgen, E.; Kemp, C.; LeMoigne, J.; Wang, L. *DRAFT Modeling, Simulation, Information Technology & Processing Roadmap Technology Area 11*; Technical Report; NASA: Washington, DC, USA : 2010.
4. Tao, F.; Cheng, J.; Qi, Q.; Zhang, M.; Zhang, H.; Sui, F. Digital twin-driven product design, manufacturing and service with big data. *Int. J. Adv. Manuf. Technol.* **2018**, *94*, 3563–3576. doi:10.1007/s00170-017-0233-1.
5. Qi, Q.; Tao, F. Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison. *IEEE Access* **2018**, *6*, 3585–3593. doi:10.1109/ACCESS.2018.2793265.
6. Boschert, S.; Rosen, R. Digital Twin—The Simulation Aspect. In *Mechatronic Futures: Challenges and Solutions for Mechatronic Systems and their Designers*; Hehenberger, P., Bradley, D., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 59–74. doi:10.1007/978-3-319-32156-1_5.
7. Negri, E.; Fumagalli, L.; Macchi, M. A Review of the Roles of Digital Twin in CPS-based Production Systems. *Procedia Manuf.* **2017**, *11*, 939–948. doi:10.1016/j.promfg.2017.07.198.
8. Enders, M.; Hoßbach, N. Dimensions of Digital Twin Applications—A Literature Review. In Proceedings of the AMCIS 2019, Cancún, Mexico, 15–17 August 2019.
9. Cimino, C.; Negri, E.; Fumagalli, L. Review of digital twin applications in manufacturing. *Comput. Ind.* **2019**, *113*, 103130. doi:10.1016/j.compind.2019.103130.

10. Barricelli, B.R.; Casiraghi, E.; Fogli, D. A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. *IEEE Access* **2019**, *7*, 167653–167671. doi:10.1109/ACCESS.2019.2953499.
11. Liu, M.; Fang, S.; Dong, H.; Xu, C. Review of digital twin about concepts, technologies, and industrial applications. *J. Manuf. Syst.* **2020**. doi:10.1016/j.jmsy.2020.06.017.
12. Lim, K.Y.H.; Zheng, P.; Chen, C.H. A State-of-the-Art Survey of Digital Twin: Techniques, Engineering Product Lifecycle Management and Business Innovation Perspectives. *J. Intell. Manuf.* **2019**. doi:10.1007/s10845-019-01512-w.
13. Bruynseels, K.; Santoni de Sio, F.; van den Hoven, J. Digital Twins in Health Care: Ethical Implications of an Emerging Engineering Paradigm. *Front. Genet.* **2018**, *9*. doi:10.3389/fgene.2018.00031.
14. Parmar, R.; Leiponen, A.; Thomas, L.D.W. Building an organizational digital twin. *Bus. Horizons* **2020**. doi:10.1016/j.bushor.2020.08.001.
15. Khajavi, S.H.; Motlagh, N.H.; Jaribion, A.; Werner, L.C.; Holmström, J. Digital Twin: Vision, Benefits, Boundaries, and Creation for Buildings. *IEEE Access* **2019**, *7*, 147406–147419. doi:10.1109/ACCESS.2019.2946515.
16. Autiosalo, J.; Vepsäläinen, J.; Viitala, R.; Tammi, K. A Feature-Based Framework for Structuring Industrial Digital Twins. *IEEE Access* **2020**, *8*, 1193–1208. doi:10.1109/ACCESS.2019.2950507.
17. Marmolejo-Saucedo, J.A.; Hurtado-Hernandez, M.; Suarez-Valdes, R. Digital Twins in Supply Chain Management: A Brief Literature Review. In *Intelligent Computing and Optimization*; Advances in Intelligent Systems and Computing; Vasant, P., Zelinka, I., Weber, G.W., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 653–661. doi:10.1007/978-3-030-33585-4_63.
18. Holler, M.; Uebernickel, F.; Brenner, W. Digital Twin Concepts in Manufacturing Industries—A Literature Review and Avenues for Further Research. In Proceedings of the 18th International Conference on Industrial Engineering (IJIE), 10–12 October 2016, Seoul, Korea; Korean Institute of Industrial Engineers, Seoul, Korea: 2016.
19. Tao, F.; Qi, Q. Make more digital twins. *Nature* **2019**, *573*, 490–491. doi:10.1038/d41586-019-02849-1.
20. Moslått, G.A.; Padovani, D.; Hansen, M.R. A Digital Twin for Lift Planning With Offshore Heave Compensated Cranes. *J. Offshore Mech. Arct. Eng.* **2021**, *143*. doi:10.1115/1.4048881.
21. Moi, T.; Cibicik, A.; Rølvåg, T. Digital twin based condition monitoring of a knuckle boom crane: An experimental study. *Eng. Fail. Anal.* **2020**, *112*, 104517. doi:10.1016/j.engfailanal.2020.104517.
22. Szpytko, J.; Duarte, Y.S. Digital Twins Model for Cranes Operating in Container Terminal. *IFAC-PapersOnLine* **2019**, *52*, 25–30. doi:10.1016/j.ifacol.2019.10.014.
23. Terkaj, W.; Gaboardi, P.; Trevisan, C.; Tolio, T.; Urgo, M. A digital factory platform for the design of roll shop plants. *CIRP J. Manuf. Sci. Technol.* **2019**, *26*, 88–93. doi:10.1016/j.cirpj.2019.04.007.
24. Autiosalo, J. Platform for industrial internet and digital twin focused education, research, and innovation: Ilmatar the overhead crane. In Proceedings of the 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), Singapore, 5–8 February 2018; pp. 241–244.
25. Sjöman, H.; Autiosalo, J.; Juhanko, J.; Kuosmanen, P.; Steinert, M. Using Low-Cost Sensors to Develop a High Precision Lifting Controller Device for an Overhead Crane—Insights and Hypotheses from Prototyping a Heavy Industrial Internet Project. *Sensors* **2018**, *18*, 3328. doi:10.3390/s18103328.
26. Chevalier, J.M.; Buckles, D.J. *Participatory Action Research : Theory and Methods for Engaged Inquiry*; Routledge, London, UK : 2019. doi:10.4324/9781351033268.
27. Valaja, A. Industry-University Innovation Collaboration: A Case of Industrial Internet Ecosystem at Aalto University. Master's Thesis, Aalto University, Espoo, Finland, 2019.
28. Wikipedia Contributors. Grounded Theory. 2020. Page Version ID: 984391878. Available online: https://en.wikipedia.org/w/index.php?title=Grounded_theory&oldid=984391878 (accessed on 20 October 2020).
29. Josifovska, K.; Yigitbas, E.; Engels, G. Reference Framework for Digital Twins within Cyber-Physical Systems. In Proceedings of the 2019 IEEE/ACM 5th International Workshop on Software Engineering for Smart Cyber-Physical Systems (SEsCPS), Montreal, QC, Canada, 28 May 2019; pp. 25–31. doi:10.1109/SEsCPS.2019.00012.
30. Grüner, S.; Pfrommer, J.; Palm, F. RESTful Industrial Communication With OPC UA. *IEEE Trans. Ind. Informatics* **2016**, *12*, 1832–1841.
31. OPC Foundation. OPC Unified Architecture Specification Part 1: Overview and Concepts Release 1.04. 2017. Available online: https://reference.opcfoundation.org/v104/Core/docs/Part1/ (accessed on 18 September 2020).
32. Aalto University. Ilmatar Open Innovation Environment. 2020. Available online: https://www.aalto.fi/en/industrial-internet-campus/ilmatar-open-innovation-environment (accessed on 10 July 2020).
33. Siemens. MindSphere. 2020. Available online: https://siemens.mindsphere.io/en (accessed on 18 November 2020).
34. Siemens. MindConnect Nano. 2020. Available online: https://www.dex.siemens.com/mindsphere/mindconnect/MindConnect-Nano (accessed on 18 November 2020).
35. Ala-Laurinaho, R. Sensor data transmission from a physical twin to a digital twin. Master's Thesis, Aalto University, Espoo, Finland, 2019.
36. Ala-Laurinaho, R.; Autiosalo, J.; Tammi, K. Open Sensor Manager for IIoT. *J. Sens. Actuator Networks* **2020**, *9*, 30. doi:10.3390/jsan9020030.
37. Ala-Laurinaho, R. Software. OSEMA: Open Sensor Manager. 2020. Available online: https://github.com/AaltoIIC/OSEMA (accessed on 21 September 2020).

38. Hietala, J. Real-Time Two-Way Data Transfer with a Digital Twin via Web Interface. Master's Thesis, Aalto University, Espoo, Finland, 2020.
39. Hietala, J.; Ala-Laurinaho, R.; Autiosalo, J.; Laaki, H. GraphQL Interface for OPC UA. In Proceedings of the 2020 IEEE International Conference on Industrial Cyber Physical Systems (ICPS), Tampere, Finland, 10–12 June 2020.
40. Hietala, J. GraphQL API for OPC UA Servers. 2020. Available online: https://github.com/AaltoIIC/OPC-UA-GraphQL-Wrapper (accessed on 5 October 2020).
41. Facebook, Inc. and GraphQL Foundation. GraphQL Current Working Draft. Available online: http://spec.graphql.org/draft/ (accessed on 29 September 2020).
42. Taelman, R.; Vander Sande, M.; Verborgh, R. GraphQL-LD: Linked data querying with GraphQL. In Proceedings of the ISWC2018, the 17th International Semantic Web Conference, Monterey, CA, USA, 8–12 October 2018; pp. 1–4.
43. Wittern, E.; Cha, A.; Laredo, J.A. Generating GraphQL-Wrappers for REST(-like) APIs. In *Web Engineering*; Lecture Notes in Computer Science; Mikkonen, T., Klamma, R., Hernández, J., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 65–83. doi:10.1007/978-3-319-91662-0_5.
44. Autiosalo, J. What is A Digital Twin? 2019. Seminar Presentation. Available online: https://youtu.be/AyCwY7ZfwKQ; https://www.aalto.fi/en/industrial-internet-campus/digitwin-demo-day-22112019 (accessed on 14 October 2020).
45. MQTT: The Standard for IoT Messaging. 2020. Available online: https://mqtt.org/ (accessed on 30 September 2020).
46. Wikipedia Contributors. Semantic Web. 2020. Available online: https://en.wikipedia.org/w/index.php?title=Semantic_Web&oldid=962469931 (accessed on 3 July 2020).
47. Contributors. Digital Twin Definition Language. 2020. Available online: https://github.com/Azure/opendigitaltwins-dtdl (accessed on 3 July 2020).
48. Valtonen, M. DT Core—A Modular Building Block for Digital Twin Data Processing. Seminar Presentation. 2019. Available online: https://youtu.be/sjQ7GDWvgCo; https://www.aalto.fi/en/industrial-internet-campus/digitwin-demo-day-22112019 (accessed on 23 September 2020).
49. Pantsar, V.; Mäkinen, H. Bring your Digital Twin to Life. Seminar Presentation. 2019. Available online: https://youtu.be/fMAtUDE9y8Y; https://www.aalto.fi/sites/g/files/flghsv161/files/2019-11/bring_your_digital_twin_to_life_by_ville_pantsar_ideal_plm_and_hannu_makinen_ideal_plm_.pdf (accessed on 22 September 2020).
50. Silvola, R. One Product Data for Integrated Business Processes. Ph.D. Thesis, University of Oulu, Oulu, Finland, 2018.
51. Tammi, K.; Takkunen, J.; Peltoranta, V.; Pantsar, V.; Autiosalo, J. DigiTwin Demo Day Morning 18.1.2019. Seminar Presentation. 2019. Available online: https://youtu.be/giZwmATMJbA; https://www.aalto.fi/en/industrial-internet-campus/digitwin-demo-day-1812019 (accessed on 24 September 2020).
52. Mattila, J. Nosturidatan Analysointi ja Visualisointi IoT-alustalla (in Finnish). Bachelor's Thesis, Aalto University, Espoo, Finland, 2020.
53. Valtonen, M.; Peltoranta, V. Flow-Based AI Methods for Hoist Brake Health Monitoring. Seminar Presentation. 2019. Available online: https://youtu.be/eT3CeiRnCSg; https://www.aalto.fi/en/industrial-internet-campus/digitwin-demo-day-22112019 (accessed on 25 September 2020).
54. Sutela, L. Digital Twin Loop Using RBS Design Automation. Seminar Presentation. 2019. Available online: https://youtu.be/8sEzUVmmDsA; https://www.aalto.fi/en/industrial-internet-campus/digitwin-demo-day-22112019 (accessed on 25 September 2020).
55. Hietala, J. Ilmatar Web app. 2020. Available online: https://github.com/AaltoIIC/Ilmatar-Web-App (accessed on 25 September 2020).
56. Hublikar, P. A Prototype of a Digital Twin with Mixed Reality and Voice User Interfaces for Controlling a Smart Industrial Crane. Master's Thesis, Aalto University, Espoo, Finland, 2020.
57. Contributors. Crane Library. 2020. Available online: https://github.com/AaltoIIC/ilmatar-python-lib (accessed on 25 July 2020).
58. Chattopadhyay, A.; Högnäsbacka, J.; Lähteenmäki, M.; Patomäki, K.; Pulkkinen, J.; Salovaara, J.; Simolin, S. Autonomous Crane for Warehouse Management—AEEproject—Aalto University Wiki. 2019. Available online: https://wiki.aalto.fi/display/AEEproject/Autonomous+crane+for+warehouse+management (accessed on 1 July 2020).
59. Lehto, T. Opiskelijoiden Tekoälykisan Voittaja Aikaansa Edellä—"Lainsäädäntö ei Vielä Salli...". 2019. Available online: https://www.tivi.fi/uutiset/opiskelijoiden-tekoalykisan-voittaja-aikaansa-edella-lainsaadanto-ei-viela-salli/a08b029e-d708-447d-b6f6-d1052f27e84c (accessed on 2 October 2020).
60. Lehto, M. Scalability of Digital Twins: Challenges and Possibilities for Efficient Implementation. Seminar Presentation. 2019. Available online: https://youtu.be/ZsG1KLiBvqk; https://www.aalto.fi/en/industrial-internet-campus/digitwin-demo-day-22112019 (accessed on 2 October 2020).
61. Tiainen, T.; Miettinen, J.; Viitala, R.; Hiekkanen, K.; Kuosmanen, P. Digital Twin and Virtual Sensor for a Rotor System. In Proceedings of the 30th International DAAAM Symposium "Intelligent Manufacturing & Automation", Zadar, Croatia, 23–26 October 2019; Number 1 in Annals of DAAAM and Proceedings; pp. 1115–1121. doi:10.2507/30th.daaam.proceedings.156.
62. Parvinen, P.; Pöyry, E.; Gustafsson, R.; Laitila, M.; Rossi, M. Advancing Data Monetization and the Creation of Data-based Business Models. *Commun. Assoc. Inf. Syst.* **2020**, *47*. doi:10.17705/1CAIS.04702.

63. Jacoby, M.; Usländer, T. Digital Twin and Internet of Things—Current Standards Landscape. *Appl. Sci.* **2020**, *10*, 6519. doi:10.3390/app10186519.
64. Azure-Samples/Digital-Twins-Explorer. 2020. Original-Date: 2020-05-29T14:35:37Z. Available online: https://github.com/Azure-Samples/digital-twins-explorer (accessed on 3 December 2020).

applied sciences

MDPI

Article

Developing a Digital Twin and Digital Thread Framework for an 'Industry 4.0' Shipyard

Toh Yen Pang [1,*], Juan D. Pelaez Restrepo [1], Chi-Tsun Cheng [1], Alim Yasin [1], Hailey Lim [1] and Miro Miletic [2]

[1] School of Engineering, RMIT University, Bundoora Campus East, Bundoora, VIC 3083, Australia; juan.pelaez.restrepo@rmit.edu.au (J.D.P.R.); ben.cheng@rmit.edu.au (C.-T.C.); s3588698@student.rmit.edu.au (A.Y.); s3776055@student.rmit.edu.au (H.L.)
[2] MEMKO Systems, Melbourne, VIC 3000, Australia; miro@memko.com.au
* Correspondence: tohyen.pang@rmit.edu.au; Tel.: +61-3-9925-6128

Abstract: This paper provides an overview of the current state-of-the-art digital twin and digital thread technology in industrial operations. Both are transformational technologies that have the advantage of improving the efficiency of current design and manufacturing. Digital twin is an important element of the Industry 4.0 digitalization process; however, the huge amount of data that are generated and collected by a digital twin offer challenges in handling, processing and storage. The paper aims to report on the development of a new framework that combines the digital twin and digital thread for better data management in order to drive innovation, improve the production process and performance and ensure continuity and traceability of information. The digital twin/thread framework incorporates behavior simulation and physical control components, in which these two components rely on the connectivity between the twin and thread for information flow and exchange to drive innovation. The twin/thread framework encompasses specifications that include organizational architecture layout, security, user access, databases and hardware and software requirements. It is envisaged that the framework will be applicable to enhancing the optimization of operational processes and traceability of information in the physical world, especially in an Industry Shipyard 4.0.

Keywords: digital twin; digital thread; framework; shipyard; industry 4.0

check for updates

Citation: Pang, T.Y.; Pelaez Restrepo, J.D.; Cheng, C.-T.; Yasin, A.; Lim, H.; Miletic, M. Developing a Digital Twin and Digital Thread Framework for an 'Industry 4.0' Shipyard. *Appl. Sci.* **2021**, *11*, 1097. https://doi.org/10.3390/app11031097

Received: 11 December 2020
Accepted: 21 January 2021
Published: 25 January 2021

1. Introduction

Digital twin (DTW) technology is the cornerstone of digital transformation, which we are currently witnessing in the new industry 4.0 revolution. DTW is accessible now more than ever and many reputable and innovative companies such as Tesla and Siemens have adopted it with varying success. Siemens [1] has integrated DTW into its three major sections of product lifecycle: product, production and performance. The virtual representation of the product is created and tested to validate performance under expected use conditions. Production is optimized through manufacturing process simulations where any sources of error or failure can be identified and prevented before proceeding to physical production. Subsequently, DTW has potential to improve performance by producing high-quality products at lowest logical cost by integrating manufacturing processes and enhancing production planning in manufacturing implementation [2].

In addition to companies in the business of manufacturing products, companies in other sectors such as the National Aeronautics and Space Administration (NASA), a pioneer of the DTW, used this technology to develop ultra-high fidelity simulation models of aerospace vehicles. These simulations enabled NASA's engineering team to predict the future performance and status of their vehicles accurately in the form of the "factors-of-safety" during design and certification phases. It also enabled mission managers to make informed decisions based on historical and real-time data to improvise possible in-flight changes to a vehicle's mission [3].

The global medical industry has been utilizing DTW to test medical devices virtually before introducing them into the physical world. For example, the Living Heart Project has adapted DTW for cardiovascular surgeons in diagnosis, education and training [4]. This project is not limited to cardiovascular surgeons but has positive implications for medical device design, clinical diagnosis and regulatory science. The fundamentals of this project involve the use of both pacemakers in live participants and virtual patients with the goal of increasing industry innovation in tackling heart diseases.

Further practical use of DTW in the medical industry relates to tailoring health care to individuals. In South Korea, DTW is being utilized in combination with Medical Artificial Intelligence to tailor healthcare plans to individual patients [5]. This, in conjunction with information on tracked health and lifestyle data from wearable devices, could eventually result in a "virtual patient." Virtual patient models allow medical personnel to perform continuous remote monitoring on patients at low-cost and provide health predictions and prescribe preventive treatments promptly. Through such interventions, South Koreans have benefitted from significant health improvements, reductions in healthcare costs and increased personal freedom in dealing with their own health.

Beyond healthcare, DTW is employed on a large scale in urban planning. For example, Virtual Singapore is a dynamic 3D city model [6] that consists of a detailed 3D map of Singapore and contains information such as texture and material representations of geometrical objects, terrain attributes and infrastructure and so forth. This 3D model is useful in virtual experimentation, virtual test-bedding, planning and decision-making and research and development.

Despite DTW being accessible now to most companies and governments, the adoption and uptake in Australian small- and medium-sized enterprise (SME) is still very slow. For most SMEs, tackling industry 4.0 problems requires a number of enabling technologies such as Product Lifecycle Management (PLM) software, enterprise resource planning (ERP) packages, the Internet of Things (IoT) and Cyber-Physical Systems (CPS), which communicate and cooperate with each other in real time. Unfortunately, it can be difficult for SMEs to integrate data into these systems when they have been developed by separate firms. Hence, the foundational knowledge, experience and potential of DTW has yet to become mainstream. There also exists a gap in understanding the requirements, applicability, security and sustainability of such technologies.

There are many studies in the field of DTW but very few studies have reported combined DTW and digital thread (DTH) technology in industrial transformations. The purpose of this paper is to report on the development of a new framework that combines the DTW and DTH for better data management in order to drive innovation, time to market and improve the production process and performance. First, we review the concept of DTW. Secondly, we consider its applicability in the entire product life cycle context. Thirdly, we describe the DTH and its entities. Fourthly, we discuss the development and integration of a DTW and DTH in a new framework and look at the necessary components for industry to embrace it. Finally, we are providing an example of combining DTW and DTH in industry 4.0 Shipyard to demonstrate how this new framework is going to work, particularly in Australian context.

2. The Digital Twin

A DTW is commonly known as a connection of data between a physical entity and its virtual representation that is made for the purpose of improving the performance of the physical part using computational simulations and techniques [7]. The concept of DTW was first introduced more than a decade ago at the University of Michigan and was further developed by Michael Grieves [8]. Grieves described the DTW as a cycle of data between three components, that is, a physical object, its virtual model and the information processing hub that links the physical object and its virtual model. Grieves envisaged this new concept as the possible foundation of PLM and a new product-manufacturing method to fulfil desired design specifications [8]. Figure 1 depicts these three components (virtual representation, information hub and physical objects) of DTW in an industry application.

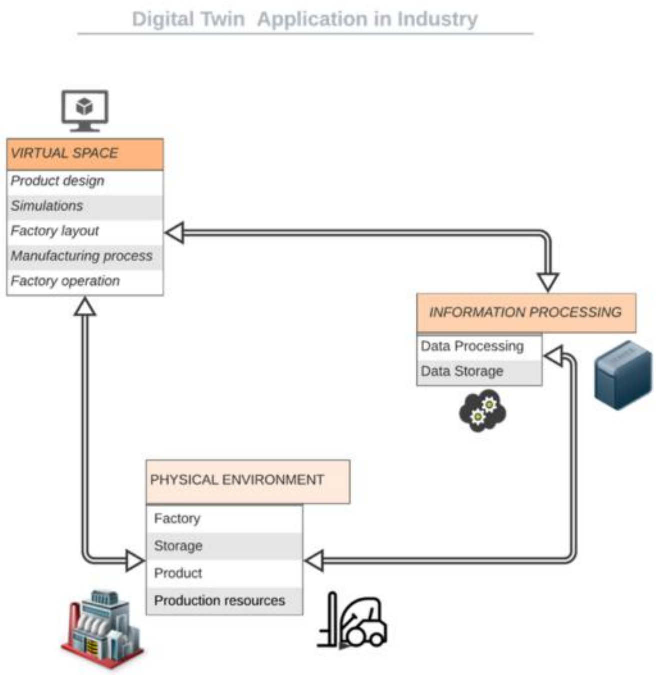


Figure 1. An example of the application of digital twin in industry.

2.1. Physical Environment

The physical environment is the basis for developing the DTW [9,10]. Generally, the objects included in most of the studies are manufactured products such as vehicles, aircraft or 3D printers. Of key importance is the fact that the DTW is not solely limited to an object itself but often considers the environment and interactions with it. If the DTW is created for the optimization of the manufacturing process, then the purpose of the DTW in the product lifecycle must be specified [7,11–13].

2.2. Virtual Space

Virtual space is the first phase of creating a DTW and incorporates a 3D model representation of the physical object, containing the geometric modelling of the physical object, the virtual workers and the virtual environment in which the product is contained. The user should model and analyze that of the 3D product in the physical space and simulate this in the virtual space, including movements of workers and the products and how they interact. The user also needs to define the attributes and properties of the product and corresponding rules of operation in the physical world and then simulate these in the virtual space. Once all these aspects have been successfully integrated into the DTW environment, the full virtual representation is considered complete.

2.3. Information Integration

Information that is collected from physical sources (from suppliers, the product itself, organizational changes) will be analyzed and integrated into the DTW during the data-integration phase. These data need to be analyzed and integrated into the DTW seamlessly. For example, a stock-taking DTW would need to understand the amount of stock left in the shop floor as physical objects and be able to translate this information to ensure up-to-date stock tracking. This is the step where the real-world data are integrated with virtual representations to create a DTW.

2.4. Current Digital Twin Application in the Industry

DTWs attract interest from different industries' operations areas such as product design, logistics, manufacturing and maintenance. Also, DTWs can be used to increase the efficiency and automation levels of the manufacturing, maintenance and after-sales service (as shown in Figure 2) [7,14,15].

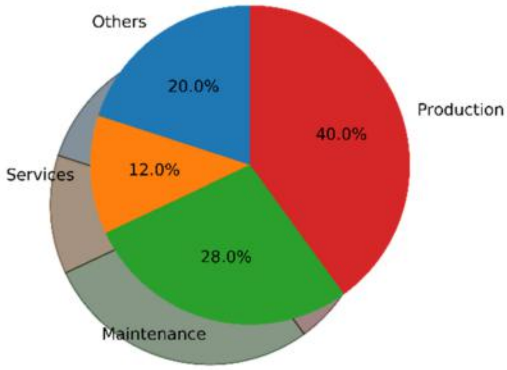


Figure 2. Application areas of Digital Twins according to Melesse et al. (reproduced from [15], Elsevier, 2020).

In the product concept, design and production phases, DTW can be a very useful tool in the manufacturing system. Studies show that DTWs have been used successfully to understand the performance and behavior of individual machines, making it easier to integrate the production line [16,17]. By leveraging the advantages of DTW, small manufacturing companies have achieved better performance in automation and adaptability to changes in customer orders or material properties such as hardness, strength and elasticity. These successes show that DTW can be used as a tool to increase efficiency in the production planning and optimization of the manufacturing implementation [16,17]. DTW has also revealed potential in the predictive maintenance area where, based on the information collected from the physical component, multi-physics simulations and data analysis are performed to predict future performance and possible future failures. These can be used to generate early warnings and to feed into the maintenance plan continuously, thereby reducing the costs of unplanned disruption. However, these kinds of applications have not yet been widely adopted and further research is needed to generalize them for wider use [9,15,18].

In the services area, such as after-sale service, DTW can be used as an information tool to provide added value to the customers by being able to produce better predictions of the future behavior and the remaining lifetime of an asset and its components. DTWs can also be used to collect useful data to drive design modification, improve product performance and improve the overall production planning cycle [15,16,19]. Despite some successful applications, the methods and tools to implement DTW in industry are still in their early stages of development and need more research. Also, many of the physical phenomena involved in the manufacturing of several products such as aircraft, vehicles and machining tools are complex and hard to simulate. Hence, these issues need more research to develop better models. Additionally, the large amount of data that can be collected by a DTW introduces new challenges in data handling, processing and storage [20,21] and hence, a framework to build DTWs should address these challenges [9,14,15,22–25].

2.5. Enabling Tools for Digital Twin

In the literature [2,26], the enabling tools for DTW can be broken down into five categories: 1. tools for controlling the physical world; 2. tools for DTW modeling; 3. tools for DTW data management; 4. tools for DTW services applications; and 5. tools for connections in a DTW environment. There are a number of commercial application platforms that have various enabling DTW technologies provided by global companies, for example, Predix (General Electric Company, Boston, MA, USA), Thingworx (PTC Inc., Boston, MA, USA), Mindsphere (Siemens, Munich, Germany), ANSYS (ANSYS Inc., Canonsburg, PA, USA), 3D Experience (Dassault Systèmes®, Vélizy-Villacoublay, France), Altair (Altair Engineering, Inc., Troy, MI, USA), Oracle (©ORACLE, Austin, TX, USA), HEXAGON (MSC Software, Newport Beach, CA, USA) and SAP (Weinheim, Germany) [26].

3. Traditional Product Lifecycle Management Approach

In the traditional PLM approaches to product development, there are many user groups and stakeholders involved in creating and sharing information during the planning, design, production and service phases (Figure 3). Hence, many documents and a large amount of data are created to capture the decisions and results of PLM activities.

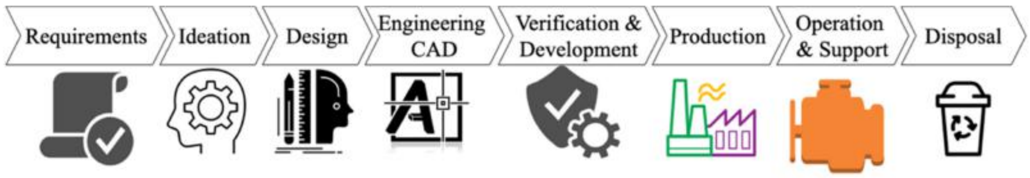


Figure 3. Traditional product life cycle management process.

Therefore, the engineers in any one team in the PLM will continually work independently by importing files locally for modification and then exporting them for storage and future use. If subsequent user groups use different data manage systems and software, the net result is that these iterations can be slow and time consuming. The overall cost required for data conversion from one part of the system to the other becomes large and reduces overall value for money.

3.1. Data Silos and Fragmented Information

For decades, organizations have optimized each product life cycle phase separate from others. Hence, highly fragmented information and knowledge exchange exists between life-cycle phases [27,28]. As a result, valuable information and knowledge is often lost and not used as context for decision-making in the transition phases and, hence, there are information gaps in the product life cycle, especially in the design-to-manufacturing and design-to-service and maintenance stages. We know that PLM is an iterative activity. Therefore, the management and exchange of information becomes crucial to ensure continuity of work flow to support innovation-based models of competitiveness and to reduce the risks of failure [27,29–31].

3.2. Digital Twins in Product Lifecycle Management

In engineering PLM, integration of DTW is a paradigm shift that can help companies set up for better processes of managing all product lifecycle stages starting from ideation, to design, testing, certification, manufacture, operations, maintenance and, finally, disposal (Figure 4) [32,33]. With a DTW, thousands of processes and modifications can be modelled for all lifecycle phases of a product. Users can test for different "what if" scenarios for changes in the design, materials, manufacturing parameters, logistics and operational conditions, among others. Furthermore, the effects of the modifications to the other phases of the life cycle can also be assessed [34].

For example, some aspects that can be achieved with DTW are a detailed recording and storage of process data from the manufacturing stages, immediate use of information from manufacturing difficulties or errors and parts defects to identify critical manufacturing steps. Also, clients can be offered customization to their needs, repair processes can be scheduled based on the knowledge of the entire product operation history throughout the product life cycle and higher machine availability, considerably lower downtimes and faster attention times following predictive maintenance of machine tools can be available [1,6].

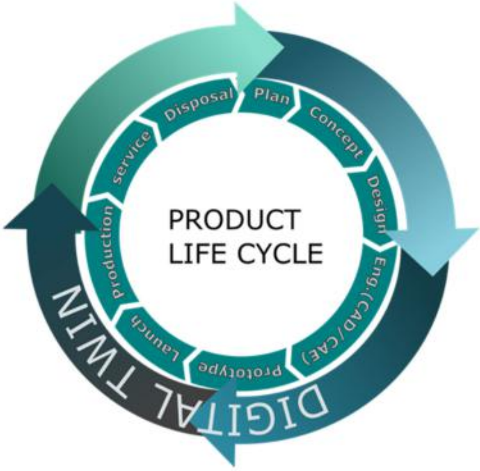


Figure 4. Integration of Digital Twin application with the Product Life Cycle management.

4. The Digital Thread

A DTH refers to a data-driven architecture that links all information generated and stored within the DTW enabling it to flow seamlessly through the entire PLM phase from invention to disposal [10,35–37]. Mies et al. [38] described the process of a DTH in the context of additive manufacturing technologies. The DTH enabled data to be integrated into one platform, allowing seamless use of and ease of access to all data. Mies et al. hypothesized that additive manufacturing processes offer ideal opportunities to apply DTH as they rely heavily on new data-driven technologies.

Siedlak et al. [39] performed a case study on a DTH that was integrated into traditional aircraft design metrics. The use of DTH enabled the necessary multidisciplinary trades to link their data through common inputs and data flows, which facilitated integrated models and design analyses. It allowed the sharing of information between usually isolated organizations to enable a more time- and cost-efficient design process.

DTH is a multi-step process that complements DTW over the entire lifecycle of the physical entity. It contains all the information necessary to generate and provide update to a DTW [35]. It relies heavily on the correct development of a framework that creates homogeneity and easy access to data through three main data chains: 1. the product innovation chain; 2. the enterprise value chain; and 3. the field and services chain (Figure 5).

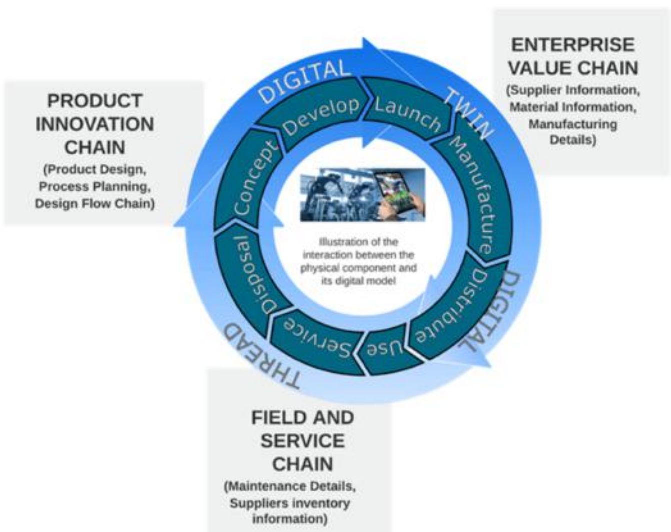


Figure 5. The concept of digital thread to complement digital twin.

4.1. Product Innovation Chain

The product innovation chain is the first step in the initialization of the DTH. This is where the lifecycle of the product is created and stored for future needs. The product designs, process planning and design flow are integrated into the thread, which outlines any suppliers and the information that were created during the first development of the physical product.

4.2. Enterprise Value Chain

The enterprise value chain is the second step in the creation of the DTH and incorporates more sophisticated details in the production of the product. This is where supplier information is integrated into the thread and on how the supplier might have produced the parts, batch numbers and so forth. Other information on the parts, including materials used and manufacturing details, would also be added. For this part of the thread, as much information can be added as the user requires. If required, all the information, including individuals who manufactured the parts, where the original materials were from and how they were obtained, can be added if this is what is required by the end-users.

4.3. Field and Service Chain

Information related to maintenance and parts is found within the field and service section of the DTH. Information that would be useful to the maintenance team and various suppliers can be seen in this section, with maintenance manuals and part availability from suppliers being incorporated into the DTH.

4.4. Key Technologies for Digital Thread

The key technologies that support implementing DTH in the three main data chains have been challenged by the difficulty in aggregating disparate data in various formats from different systems and organizations throughout the product lifecycle [36]. There exist commercial software tools that support inter-operability and enable the DTH applications. For example, the ModelCenter (Phoenix Integration, Blacksburg, VA, USA), TeamCenter (Siemens, Plano, TX, USA), ThingWorx (PTC Inc.), 3DExperience (Dassault Systèmes®), Aras Innovator® (Aras, Andover, MA, USA) and Autodesk Fusion Lifecycle (Autodesk Inc., San Rafael, CA, USA) are various commercial software tools for managing centralized data storage and the integration of simulation models for optimizing product and system designs [40–42].

5. New Digital Twin and Digital Thread Framework Development

The importance of DTW and DTH is highlighted by academe and industry due to its virtual/real-world integration [9]. As DTWs can integrate data collected from physical models with data from computational models and processes with advanced prediction methods, the results can be used to improve the performance of the existing product or to produce improved versions in the future [7]. Also, product design, assembly, production planning and workspace layouts have been found to be potential fields for twin/thread framework application [17,43].

The development of a new DTW and DTH framework (hereafter, the twin/thread framework) is an integration of DTW and DTH step and often requires more resources than when building a DTW for the first time. The twin-thread framework has multi-layered stages (Figure 6), which require the developer to follow a loop-style iterative approach to develop it.

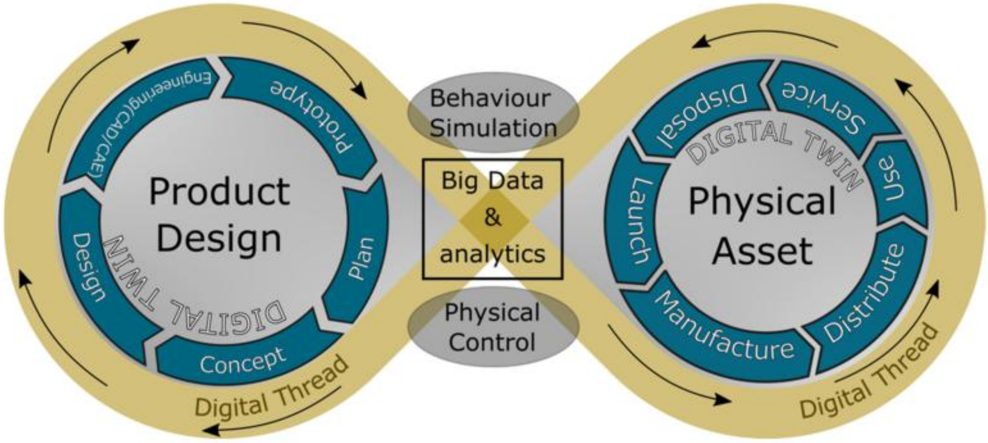


Figure 6. Digital Twin and Digital Thread framework for efficient product data management.

The new twin/thread framework comprises product design and physical asset components that are building blocks for establishing a centralized product data management (PDM) system. A PDM system will ensure the inter-operability of services and platforms involved in a project and help to standardize file formats, adopt common data storage and representation approaches and impose version control on data files across platforms. In addition, the PDM system will not only save time for engineers and designers in importing files from one platform and exporting them into another but also allows them to communicate and collaborate constantly with other stakeholders (i.e., a non-linear style approach) via a unified and consistent data representation framework, with the aim to delivering relevant data to the right person at the right time and in real time.

The advantage of the new twin/thread framework is that the users can use DTW to set up virtual models to test out scenarios to investigate where problems might have occurred and help them to predict what they might do to rectify the problems. DTH is an added benefit where it enables all stakeholders to effectively communicate and share big data bi-directionally up and down stream throughout the entire product life cycle.

5.1. Integration of a Model-Based Systems Engineering (MBSE) Approach to Support PLM

Given the increasing model-driven data across many industries, a new Model-Based System Engineering (MBSE) approach was introduced. MBSE uses a unified platform to support the requirements of design, analysis, verification, production and maintenance within the entire PLM activities [29]. MBSE aims to use a models-oriented approach (instead of document-based approach) to support the exchange of information. Figure 7 provides an overview of the common tiers of MBSE architecture. The lowest tier in the architecture contains data that are to be accessed and potentially used for analysis. Systems within the middle and top tiers provide functions and services that manage the translation and/or transaction of data between different organizations [36].

The decision-makers can also use MBSE to manage risks by defining proactive and reactive resilience strategies and contingency plans using the historical and real-time disruption data analytics to ensure business continuity [44].

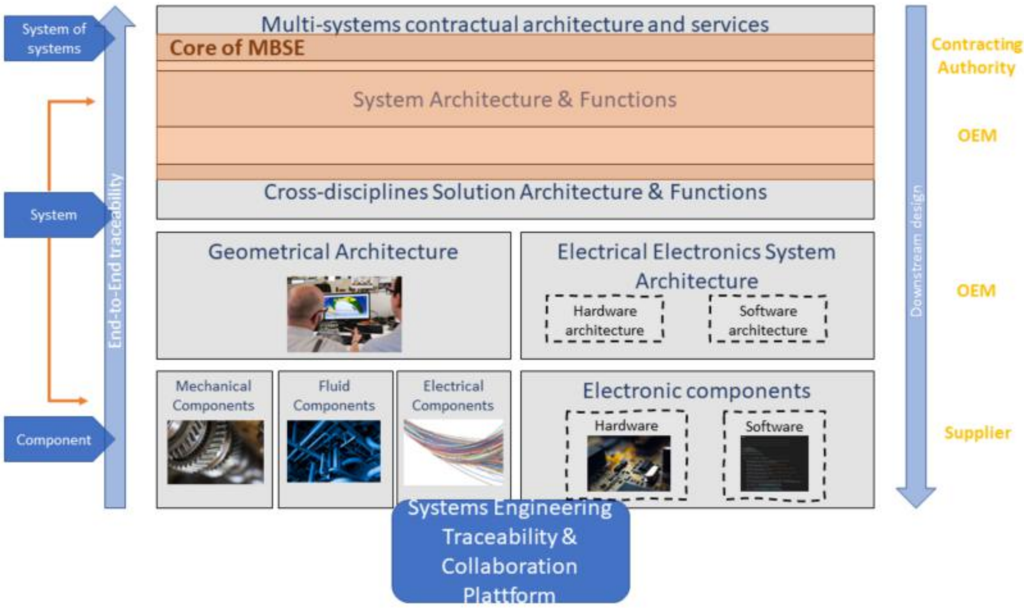


Figure 7. Model-Based Systems Engineering framework to support the consistency of exchanging model-data.

5.2. Behaviour Simulation

An operation of a process is required for the behavior simulation step to simulate a physical product in a virtual space. Key functions of the physical model will be simulated and the response of the virtual product will be examined. For example, in a stock-take model, the virtual model could be simulated to represent a real-life scenario of lost stock, the virtual model would then be required to find the supplier and order new stock to replenish the resources automatically, keeping the flow and function of overall product. Behavior simulation needs inputs from the DTH with respect to supplier information to be integrated into the DTW. Once the behavior is simulated virtually, the system can move to the physical control and complete the twin/thread cycle.

5.3. Physical Control

Physical control is the last stage in the twin/thread framework and involves controlling and changing the physical system. The physical control brings the other steps together and produces a fully functioning DTW that can change and interact with the physical model. By incorporating sensing and controlling systems and linking them with the communication infrastructure, the physical model will be able to be manipulated and changed within a virtual space. The behavior and structure of the physical world can be controlled manually or automatically through the DTW and real-world changes can be analyzed and optimized through simulations. After the physical control has been executed, the DTW will update instantly to simulate the new physical model. For example, for a stock-take delivery on-time set-up, the use of sensors would identify a low stock levels of a product and the product would be ordered through the supplier information based in DTH and the DTW would be updated with the amount of stock. Once delivered, the stock would then revert to 'normal' supply levels and the DTW will need to be updated immediately to reflect this change.

Once physical control is completed, the next iteration of the cycle begins and the DTW will need to be constantly updated in order to keep up with the workforce and the demanding needs of the new industry 4.0.

The twin/thread framework also encompasses different aspects including organizational architecture layout, security, user access, data storage and hardware and software requirements, which will be addressed in the following sections.

5.4. Organisational Architectures

First, the organizational architecture needs to be developed in the system. This may be set up by the supplier of the software or can be set up in-house depending on the users' needs. This includes the organization set-up, logos and context behind the DTW before starting the process, setting up a clear outline of what the organization needs and the needs of the users.

5.5. Data Storage Requirements

The software requirements for the twin/thread framework also need to be established for the data to be easily managed and imported into the various systems. Ideally, the software would allow for all the functions required in the DTW including 3D modelling, product design chain flow, manufacturing details and service information. Whichever software is chosen by the user should also include a service agreement with that company to ensure any complications and issues can be resolved, enabling maximum efficiency and use of the software.

A large volume of data will be collected from a variety of sources during the entire PLM process. These data can be classified into three sets: 1. structured (i.e., data with specific formats such as digits, symbols, tables, etc.); 2. semi-structured (e.g., trees, graphs, XML documents, etc.); and 3. unstructured (e.g., texts, audios, videos, images, etc.) [45] These data need to be stored in databases for further processing, analysis and decision-making. Big data storage technologies, such as distributed file storage (DFS), standard Structured Query Language (SQL) databases, NoSQL database, NewSQL database and cloud storage, can be applied according to the nature of the data [26,45].

The DTW model can be updated continuously with the newest data stored in the database via SQL queries or online application programming interface (API). Interactive dashboards and other visualization tools, such as AR/VR goggles, can extract and consume data using the same mechanism.

5.6. Hardware Requirements

The hardware requirements for the software also need to be established before developing the twin/thread framework. These requirements are based on what software the users will be running for the DTW (examples of software include 3DExprience) and the type of activities they will be undertaking with the software. For CPU-exhaustive tasks (such as design tools or CAD creation), premium hardware is needed to run the required software. As there are a number of companies that offer the enabling software and technologies for DTW, users are recommended to refer to the vendor's certified hardware specifications. For example, Dassault Systèmes' has its specific certification process for workstations and laptops from various manufacturers, models, operating systems, graphic cards and drivers. This is to ensure reliable operation and seamless integration of the DTW enabling software and removes any hardware issues in running the software. It is also recommended that the hardware be upgraded periodically to ensure smooth operating and functionality for all users. By investing time in development, the twin/thread framework will run effectively by eliminating compatibility and scalability issues. Without investing time in the framework, users might experience poor software instability run time and will lack productivity due to a non-sustainable software environment in the long-term use of the twin/thread framework.

5.7. Cyber Security Framework

The next vital step is to set up and control cybersecurity for the twin/thread framework to ensure cyber resilience (Figure 8). The cybersecurity protocol contains three essential elements: 1. robust policies to maintain safeguard; 2. technologies that comply with security control; and 3. training of staff to support organizational awareness [46]. Data security could be industry-specific and some industries might require more rigorous security measures than others. A measure that would ensure the safety of the information in the twin/thread framework would be the implementation of ISO27001 [47]. ISO27001 is an international security standard developed to provide a model for establishing, implementing, operating, monitoring, reviewing, maintaining and improving an information security management system. These security measures could be implemented to all users who have access to the DTW on the server. Additional training is recommended to all users to ensure the utmost safety of the organizations and the information stored within the twin/thread framework. This is an integral stage in the framework's development, as this is what protects both the users and the suppliers from potential danger and IT crime [46].

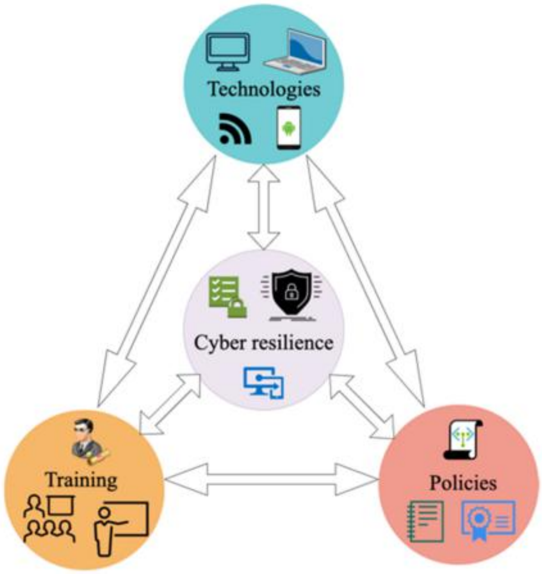


Figure 8. Cybersecurity framework for the digital twin/thread system.

Identifying correct user access and the creation of an identity and access management protocol (IAM) for the user are the next stages of the framework's development. This involves setting up correct access and roles for the right users, ensuring that only the information and recourses that are needed by that user is accessed [48,49]. User authorization needs a further authentication step to ensure the security of the data. This could be achieved through adapting strong or multi-factor authentication options such as the use of security questions or through email authorizations [48].

5.8. Proposed Architechture of Enabling Digital Twin/Thread Application

Despite DTH spanning the entire product life cycle, digital data continuity from the design to maintenance stage, as well as between Original Equipment Manufacturers (OEMs) and suppliers is limited. 'Discontinuity' of digital data and fragmentation of supply chain information might be the result from the use of many CAD software and/or PLM systems by OEMs, cyber security and data sharing control requirements and the lack of the required technology and digital skills among OEMs and suppliers.

A new enabling framework is, therefore, needed to link all information within the DTW to flow seamlessly through the entire product life cycle to support downstream processes in real-time and to address the challenges from design to manufacturing transition. The new enabling framework should have sufficient functionality, scalability and connectivity with customers and suppliers to ensure digital continuity and be easily integrated into the twin/thread framework.

In order to achieve digital continuity in the entire PLM, a platform and a set of software applications dedicated specifically to engineering design, verification and manufacturing are required. As noted in the literature [50,51], standardized design software, databases, tools and processes are a key to success for big and complex projects that involve many stakeholders from many countries to ensure digital continuity and traceability without causing costly mistakes and delay. Figure 9 provides an overview of the proposed architecture of a twin/thread system, which comprises organization/technical specifications, associated interface tools, PLM components, data analytics and the operation of the model-oriented MBSE approach. Each aspect of the proposed architecture is discussed in the following sub-sections.

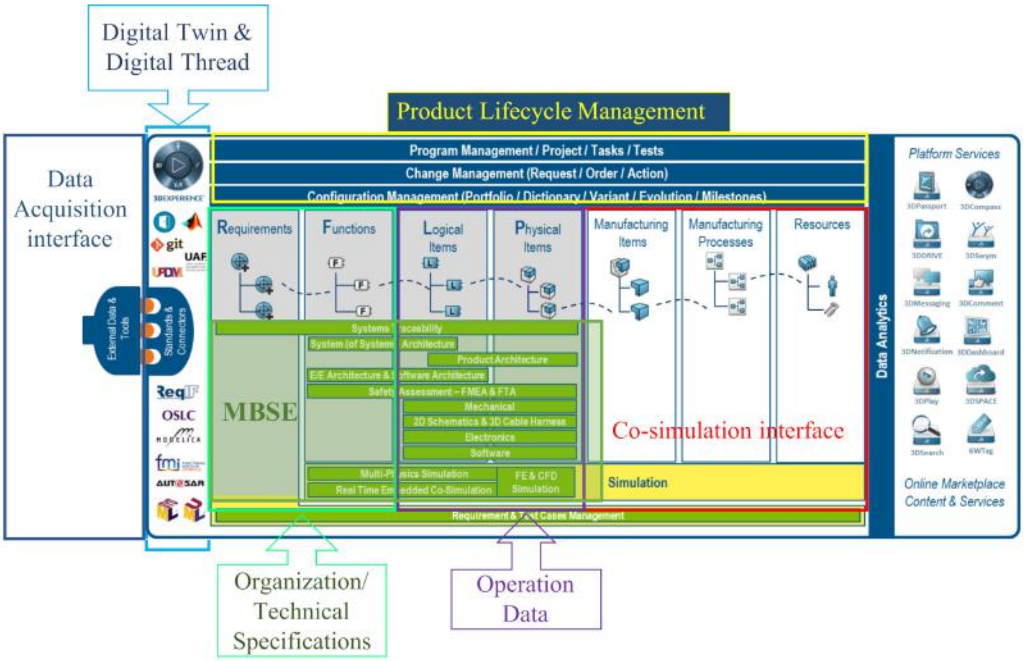


Figure 9. Proposed architecture for enabling digital twin/thread application to enhance digital continuity and traceability.

5.8.1. Digital Twin and Thread Application Suites

The top section of the framework includes a database, an application server and thick client (i.e., software such as 3DExperience). The application server provides the interface between the database and access to internal and external clients [36]. The database contains interdisciplinary models, for example, CAD models, functional models and simulation models. Each of these models are created during the engineering process of a DTW using specific tools. The connectors such as 1. Open Services for Lifecycle Collaboration (OSLC) links, which establish traceability and analyze relationships between the requirements, functions, resources, manufacturing and processes, 2. AUTomotive Open System ARchitecture (AUTOSAR)—a standard for system specification and exchange that helps to improve the reusability of vehicle software architectures, and 3. Unified Profile for Department of Defense Architecture Framework (DoDAF) and the UK Ministry of Defence Architecture Framework (MODAF) (UPDM)—a common software language to describe defense architectures, are used to connect data and achieve the DTH across domains, applications, organizations, systems and systems-of-systems. The DTW and DTH are connected to the PLM data repository via the data acquisition interface.

5.8.2. Product Lifecyle Management Components

Users can employ the PLM features to configure the collaborative creation and management dissemination of information related to product. These features allow users from different locations to work concurrently in real time on the same data, via a simple web connection to the twin/thread application suites. The integration of such features within the twin/thread suites allows users to optimize the change management processes as well as minimize the impact on every stage of the lifecycle [52].

5.8.3. Model-Based Systems Engineering (MBSE)

MBSE provides a common guideline on the management concept, system-to-system architecture and operational scenarios to promote concurrent model development and enhance re-usability of model data. It aggregates the model data from engineering and manufacturing items and processes or from different organizations in the supply chain. With MBSE, users can employ modelling and simulation data to create DTW of the physical assets in each step of the lifecycle journey. Then, the DTH will link its corresponding DTW to the design of the physical systems to ensure traceability links [52]. See Section 5.1 for details.

5.8.4. Data Acquisition Interface

The data acquisition interface will capture and store data collected by sensors and operational data from the real world. Through this interface, sensor and operational data can be transferred to DTW and this, subsequently, allows users to perform dynamic behavior simulation in parallel with the real-world data. Technologies that can implement a data acquisition interface are, inter alia, Predix (General Electric Company), Thingworx (PTC Inc.), Mindsphere (Siemens) and 3D Experience (Dassault Systèmes®) [26].

5.8.5. Organization and Technical Data

These data contain information about the physical asset itself. All documentations (e.g., requirements, specifications, design layouts, service manuals, maintenance reports etc.) that are generated by all stakeholders throughout the entire product life cycle can be stored here [52].

5.8.6. Operational Data

Real-world operational data can also be stored here using the data acquisition interface [53], such as: 1. sensor data, which is continuously streamed and recorded the current operation of an asset; 2. control data, which determines the current status of the real component; and 3. Radio-Frequency Identification (RFID) scanner data, which capture the current physical location of physical assets.

5.8.7. Co-Simulation Interface

The co-simulation interface can be used to simulate the flow of the entire production system and manufacturing processes in the real world [53]. For example, a user can utilize a factory layout program in the application suite to create a DTW of an existing physical factory floor and then use the factory flow simulation interface to create the factory flow process, starting from the supply of raw materials to the final dispatch of end products. The user can begin the simulation by choosing a start location, the required resources (e.g., 3D objects used in simulation, raw materials, worker manikins, etc.) and manufacturing processes (e.g., conveyor belts, numerical control machines, robotic arms, etc.) for the designated tasks. While the simulation is running, the current state, utilization percentage, current capacity and total of activities completed can be tracked. A system performance monitor in the simulation interface can be used to display live information for the whole factory floor and all resources. The live information, which includes utilization, total activities completed, average bottleneck of resources and current operational state of machinery, can provide unique and important support to customers and shareholders in implementing strategic planning and optimization.

5.8.8. Big Data and Analytics

There is a necessity for a platform for reliable 'big data' storage and to perform data analytics for decision-making. A large amount of data is generated and processed at any stage of the product lifecycle [54]. Large datasets can also come from various sources (e.g., computers, mobile devices, sensing technologies) [55,56]. Data analytics provides the capacity to analyze large and complex datasets and project/process managers can gain greater insight to make informed decisions and implement actions by searching, discovering and processing patterns in big data [55]. When a product is manufactured, all relevant data, such as status data from machines or energy consumption data from manufacturing systems, are stored and accessible in the DTW via the data acquisition interface. As a result, energy consumption optimization and better operational efficiency can be achieved. Such data also provide actionable insights for future decision-making.

5.9. Intellectual Property (IP)

In a globalized environment where innovation is crucial, the main competitive advantage of organizations lies in the development of new ideas and intellectual property (IP). Throughout the phases in a product's lifecycle, many change iterations (e.g., changes in customer demands or amendments in design and optimization) and the exchange of many highly sensitive information (e.g., IP products and services or personal information) will take place between various user-groups and stakeholders [57]. Historically, organizations faced a lack of integrated systems to manage their IP and heavy reliance on spreadsheets/manual documents. Thus, managing IP protection raises numerous challenges for organizations. The following sections elaborate on how the proposed approach can help to ensure IP continuity and protection.

5.9.1. Intellectual Property Continuity

Traditionally, organizations use the "throw it over the wall" approach, where different teams work in insolation from each other. Once a task is completed, they will hand over documents and 3D models to the next team. This approach does not address data silo issues and information that is often lost or lacks traceability [10]. The proposed twin/thread framework can play a significant role in modern product development and management. It provides a single, shared PDM platform to connect various user-groups and stakeholders throughout the entire product lifecycle from concept to disposal. The PDM platform will: 1. allow users to have easy, quick and secure access to data in a central repository during the product design process and 2. enable users to support product development and management processes by sharing, updating and controlling the way users create, modify and monitor the flow of product-related information. Such processes occur during entire the product lifecycle and each stage involves dynamic interactions between entities that use the available information to generate new information and IPs and share them further [57,58]. As such, the proposed twin/thread framework will transform the way organizations manage their information and IP more efficiently by harmonizing all sources and types of data (of different formats, stored using different means and in different locations) to ensure digital continuity and traceability (Figure 10).

Figure 10. Management of intellectual property to ensure its continuity and filling the missing gaps.

As IP management maturity increases, companies can identify gaps related to engineering design, manufacturing planning, steps of a production process and service and maintenance over the lifecycle. With the ease of tracible information and knowledge, organizations can fill in the missing gaps for generating real growth possibilities [59].

5.9.2. Intellectual Property Security and Protection

One issue is how to protect IP effectively from loss, leak and theft. Through the adoption of a model oriented MBSE approach with the twin/thread framework and with proper cyber-security measures in place, organizations can provide segregated access to internal and external clients (e.g., OEMs and suppliers). In this regard, a number of appropriate organizational and technical concepts to exchange, manage and control access to information securely will be considered [49,60]:

1. Role-based access control. This allows organizations to manage a user's role and access to documents and directories. Once the user has authenticated his or her personal information against the system using a username and password, the system will grant access based on the defined role,
2. Digital watermarks. These provide a unique identification of origin for a document that can easily be traced when it is made accessible to internal and external stakeholders,
3. Data Leakage Prevention (DLP). This blocks extracting files by external, non-authorized devices such as USB sticks and keep track of e-mail traffic and the flow of information and its use, with whom it is shared and what actions are applied to it [57], and
4. Enterprise Rights Management (ERM). This integrates the know-how in suitable CAD and non-CAD (e.g., pdf-documents, MS Office documents) templates, which are encrypted using ERM-templates during the creation process and can be decrypted only after authentication against the decryption key received from the ERM server [60].

Depending on the business model and the needs of the organization, commercial software providers (as identified in Section 2.5) can provide consultation, implementation, integration, hosting and training services for potential control of access to information from PDM and PLM platforms, and from a shared folder to companies with different scales. This secure access can ensure the protection of IP and other proprietary lifecycle data [36], for example, the IP and the design of the product being fabricated, batch 'Bill of Materials' components and any processes being developed to fabricate the product [36,61].

6. Industry 4.0 Shipyard

The differences between implementing a DTW in the manufacturing and maritime domains have been recently studied [56]. The study showed that very few implementation frameworks for the maritime domain have been developed but found one promising framework with the basic requirements for a DTH solution. The study concluded that both domains are developing open platforms for DTW implementation and present some useful real-world implementation examples of DTW [56].

DTW has also been proposed as a natural step from MBSE, with great possibilities of improvement in the production of highly complex products such as cruise ships. Some of the advantages highlighted are the ease in collaboration amongst all teams involved in the process of ship design [62–65]. Also, the possibility to access information and manage it efficiently using an advanced interface could help develop efficient maintenance and training programs that, in time, can lead to higher operational performance levels [66,67].

Additionally, DTH has been identified as a different way from the traditional 2D drawings for shipbuilders to design and build their ships faster and better. DTH offers shipbuilders the possibility of having their employees and suppliers connected to and synchronized with their shipyard, production planning, customer orders and requirements, 3D models and every aspect of design [68].

However, the current DTW models applied in the ship building industry show that only some of the components of the ship are being represented in the DTW, which is understandable due to the considerable number of sub-assets included in a modern ocean vessel. Including such a huge number of parts and their properties, interactions and performances in a model imposes great challenges. As DTH technology continues to mature, it will help the industry improve several aspects of their production processes through collaboration and constant communication of information [69].

6.1. Proposing Digital a Twin/Thread Framework in Australian Shipyard

As a result of the progressive implementation of smart and autonomous systems of Industry 4.0, the shipbuilding industry has developed a new, radical paradigm in its manufacturing systems by integrating automated tools and processes, creating new demands for more lean production processes, while increasing production efficiency, improving ship safety and reducing environmental impacts.

Furthermore, in a very complex shipyard site that contains large areas for fabrication of the ships, dry docks, slipways, warehouses, painting facilities and so forth, there exists many moving goods and many parts may look alike during the entire ship building life cycle. Hence, there is a need for ship operators to develop a relatively energy efficient way of moving goods and to accurately identify and trace moving goods to minimize impacts and to improve productivity and safety in the shipbuilding process.

In developing Shipyard 4.0, we believe a right framework is required to assist in designing a virtual work environment using highly detailed DTW, which could optimize the entire shipbuilding process by delivering the right information at the right time to avoid mistakes and increase productivity.

The concept of the DTH is typically defined as the flow of information that informs how a product moves through its design and production lifecycle (Figure 11). The implementation of DTH allows the monitoring of production in shipyard facilities and of the suppliers' production in their own plants. This provides greater product and process visibility to the ship builders, as well as greater transparency for the customer throughout the building process.

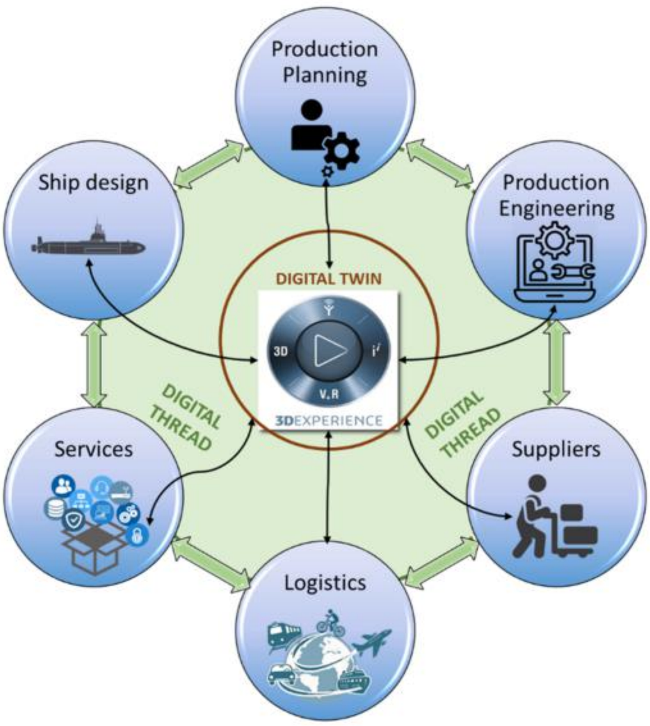


Figure 11. Digital Twin and thread implementation scheme for a shipyard.

A DTW of the shipyard will improve the efficiency of the factory flow when data can be extracted via behavior simulations of an operation process, such as machining time. Extracted data can be connected via DTH and fed into the DTW to identify bottlenecks. Furthermore, having a twin/thread framework where the DTW resides offers benefits in time management in scheduling and delivery. This relates to the inclusion of supply-chain data in the DTW. A fully integrated supply chain allows users to access the full spectrum of information available. A twin/thread framework could improve decision-making thanks to its single source of information.

In addition to improving manufacturing and design stages, DTW and DTH could enhance managerial decision-making processes. Provided that a true DTW of the shipyard is created, the information generated from all areas would be conducive to optimizing and achieving key performance indicators. Additionally, if the supply chain was integrated into the DTW, this information would give management information on what to expect, potential future issues and time to adjust to unforeseen circumstances.

Ships are normally built to last for up to thirty or more years. Therefore, it is important to ensure the continuity and traceability of design-to-service and design-to-maintenance information until their final dismantling. After its construction ships will continue to operate in the seas and will have impacts on the environment throughout their operational lives. The use of the twin/thread framework with the integrated MBSE and big data will be able to help providing a way of more effectively dealing with environmental and other issues.

According to the Australian Naval Group's SEA 1000 program [70], a total of 12 submarines will be built and all expected to be in operation by the mid-2050s. When considering future design aspects, over the next 30 years, the DTW offers the opportunity to test and reiterate designs via virtual testing, such as thermal and structural analysis, for improvement through its feedback loop processes. Legacy, historical and real-time data (maintained history, sensors data, test results etc.) connected via DTH through the physical ship can be subsequently fed back into the design process and used to improve design in case there are unforeseen circumstances or realized areas of improvement.

6.2. Roadmap for the Implementation of the Proposed Twin/Thread Framework

The implementation of the twin/thread framework can be challenging for an organization. A clear understanding of the framework and careful planning are essential to deploy its applications effectively to meet the organization's requirements and needs and to prevent costly mistakes. There are a few global companies that provide 'out-of-the-box' software applications and PLM solution suites for both DTW and DTH, including PTC Inc., Siemens, ANSYS Inc., Dassault Systèmes® and Autodesk Inc. For organizations interested in implementing the twin/thread suites as a mean to improve efficiency, software providers would normally offer consulting, implementing and support services that align with the customers' business requirements. While every implementation journey is unique, businesses can obtain the best results by following an industry 'best practice' and the methodology roadmap shown in Figure 12. The common phases for implementing out-of-the-box twin/thread solution are divided into: 1. access and definition; 2. design and build; and 3. deployment and support.

The very first step is for businesses identifying the requirements and needs for the twin/thread suites within the enterprise. A comprehensive understanding of an organization's own business processes and requirements can provide insight into how to set up the necessary organizational architectures to ensure the seamless flow of information outlined in Section 5.4. Once the requirements have been identified, it is time to engage a potential software provider and system integrator to put the plan into action by identifying the hardware, software and data storage requirements (Section 5.5 and Section 5.6) and nominating the project team to champion the roles. Prior to full roll-out, it is important to design and build the distinct architectures of the twin/thread applications that are aligned to the organization's requirements. The proposed architecture (Section 5.8) can be used to guide implementation.

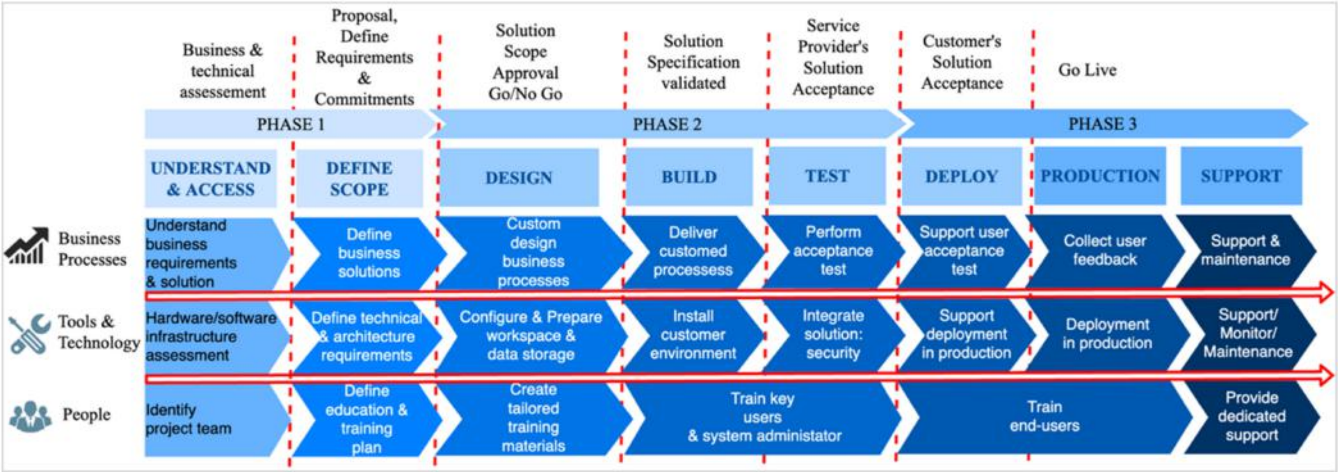


Figure 12. Design, build and implementation of the 'out-of-the-box' digital twin/thread product roadmap (modified from [71]).

While the twin/thread suites have effective hardware and software security processes, it is critical for organizations to consider the additional security measures outlined in Sections 5.7 and 5.9.2, especially in cases that involve IPs and new innovations when the twin is replicating their physical counterparts throughout the entire lifecycle [72]. Early participation from executive leadership and a well-trained and educated workforce in the twin/thread suites is a key attribute to ensure successful implementation in an organization. Finally, regular database-integrity checks and maintenance need to be considered before the application goes live and beyond to ensure that any problems are detected and administrators can either restore from a backup or conduct repair options.

6.3. Operation and Sustainment of Twin/Thread Framework in Australia Shipyard

With the adoption of the twin/thread framework, the shipyard industry can utilize DTW to transform the whole production lifecycle to ensure sustainability and to improve the performance of future programs [10,55]. For example, design engineers can leverage MBSE to work together with manufacturing engineers to create 3D models and simulations that link to real-time visualization for digital and physical production processes and instructions throughout the entire product value chain. The DTH will provide a platform to aggerate big data from disparate systems throughout the product lifecycle into actionable information through data analytics. With this deep insight from diagnostic analytics, descriptive analytics and predictive analytics, engineers, managerial teams and technician can use the data to support decision-making [55].

The twin/thread framework has been proposed and, in order to make it work, contracting authorities in the shipyard industry need to have the necessary hardware and software systems to facilitate multi-OEM participation in the DTH to ensure the connectivity of data. The sustaining of twin/thread frameworks will depend on continuing digital transformation, the endorsement of standardize tools and data exchange and better understanding of and agreements for upstream lifecycle functions to accommodate needs in downstream functions [10].

7. Conclusions

DTW and DTH are two promising technologies that will allow the manufacturing industry to optimize of their operational processes and traceability of information in the physical and virtual worlds. However, from the literature reviewed in this article, it can be concluded that these technologies are still in their early stages and further research related to implementation is needed, especially in framework development and in data processing, storage and security.

At present, existing frameworks can perform only limited aspects of what a true DTW and DTH should be able to achieve. While a DTW is designed to include the entire lifecycle of a physical part from design to use and then disposal, existing frameworks are largely focused on the design and creation stages only. Though some papers have referred to PLM in relation to DTH, which to ensure the connectivity of data silos and isolated information and elements to improve communication and collaboration, the existing DTH technology that integrates seamlessly with DTW has yet to be successfully implemented.

The proposed twin/thread framework, which uses DTW to represent the enterprise chains (i.e., product innovation chain, enterprise value chain and asset chain) and uses DTH to connect the enterprise data together to create digital continuity and accessibility. The advantage of the new twin/thread framework is that the users can use DTW to set up virtual models to simulate possible scenarios to predict future performance and the possible future failures. DTH is an added benefit where it enables all stakeholders to effectively communicate and share big data bi-directionally up and down stream throughout the entire product life cycle.

In order to adopt the twin/thread framework, OEMs need to define and adopt suitable technologies for product, process and resource modelling and validation, then maintain a digital repository for the deposition of the numerous products, processes and resources information within a single platform, of which the Model based System Engineering (MBSE) approach was introduced. The MBSE approach allows user-groups and stakeholders to collaborate on a unified system, where they can share data, perform simulation and visualization of a highly detailed model of a future physical product and exchange information in the form of models instead of document.

This will open avenues for accurate identification and easier traceability of information that will lead to improved efficiency and productivity. More significant is the possibility of iterative designs through feedback processes, which can shorten production lead times. This feedback is made possible through the DTH connecting the physical environment and DTW and is created through data extractions from both the physical and digital worlds. The same information that improves future design is used for management decisions.

In the context of the shipyard, the benefits of integrating a twin/thread framework into the established shipyard process span improved productivity and performance. Design engineers can leverage on the DTW to test and reiterate designs via virtual testing, such as thermal and structural analysis, for improvement through its feedback loop processes. The DTH will provide a platform to aggerate big data from multiple sources, such as maintained history, sensors data, test results and so forth, throughout the product lifecycle into actionable information through data analytics to improve the performance of future programs.

Author Contributions: Conceptualization, T.Y.P., C.-T.C., M.M. and A.Y.; Methodology, J.D.P.R., A.Y. and H.L.; Software, C.-T.C. and M.M.; Investigation, T.Y.P., A.Y. and H.L.; Resources, J.D.P.R., A.Y. and H.L.; Writing-Original Draft Preparation, T.Y.P., C.-T.C., J.D.P.R., A.Y. and H.L.; Writing—Review & Editing, T.Y.P., C.-T.C. and M.M.; Visualization, J.D.P.R., A.Y. and H.L.; Supervision, T.Y.P., C.-T.C. and M.M.; Project Administration, A.Y., H.L. and J.D.P.R.; Funding Acquisition, T.Y.P. and C.-T.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Defence Science Institute (DSI), grant number CR-0032.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: The authors thank the students who participated in this research. We also acknowledge the contributions of the staff of MEMKO Systems Pty Ltd.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Barrios, P.; Eynard, B.; Danjou, C. Towards a Digital Thread Between Industrial Internet of Things and Product Lifecycle Management: Experimental Work for Prototype Implementation. In Proceedings of the 16th IFIP International Conference on Product Lifecycle Management (PLM 2019), Moscow, Russia, 8–12 July 2019; Springer: Cham, Switzerland, 2019; pp. 273–282.
2. Qi, Q.; Tao, F.; Zuo, Y.; Zhao, D. Digital Twin Service towards Smart Manufacturing. *Procedia CIRP* **2018**, *72*, 237–242. [CrossRef]
3. Glaessgen, E.; Stargel, D. The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles. In Proceedings of the 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Honolulu, HI, USA, 23–26 April 2012. [CrossRef]
4. Scoles, S. A Digital Twin of Your Body Could Become a Critical Part of Your Health Care. Available online: https://slate.com/technology/2016/02/dassaults-living-heart-project-and-the-future-of-digital-twins-in-health-care.html (accessed on 20 June 2020).
5. Shin, S.Y. Current status and future direction of digital health in Korea. *Korean J. Physiol. Pharmacol.* **2019**, *23*, 311–315. [CrossRef] [PubMed]
6. Goto, S.; Yoshie, O.; Fujimura, S. Empirical Study of Multi-Party Workshop Facilitation in Strategy Planning Phase for Product Lifecycle Management System. In Proceedings of the 16th IFIP International Conference on Product Lifecycle Management (PLM 2019), Moscow, Russia, 8–12 July 2019; Springer: Cham, Switzerland, 2019; pp. 82–93.
7. Jones, D.; Snider, C.; Nassehi, A.; Yon, J.; Hicks, B. Characterising the Digital Twin: A systematic literature review. *CIRP J. Manuf. Sci. Technol.* **2020**, *29*, 36–52. [CrossRef]
8. Grieves, M. Digital twin: Manufacturing excellence through virtual factory replication. *White Pap.* **2014**, *1*, 1–7.
9. Tao, F.; Zhang, H.; Liu, A.; Nee, A.Y.C. Digital Twin in Industry: State-of-the-Art. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2405–2415. [CrossRef]
10. Leiva, C. Demystifying the Digital Thread and Digital Twin Concepts. *Ind. Week*. 2016. Available online: https://www.industryweek.com/technology-and-iiot/systems-integration/article/22007865/demystifying-the-digital-thread-and-digital-twin-concepts (accessed on 18 August 2020).
11. Hofmann, T. Integrating Nature, People, and Technology to Tackle the Global Agri-Food Challenge. *J. Agric. Food Chem* **2017**, *65*, 4007–4008. [CrossRef]
12. Mohammadi, A.; Jahromi, M.G.; Khademi, H.; Alighanbari, A.; Hamzavi, B.; Ghanizadeh, M.; Horriat, H.; Khabiri, M.M.; Jahromi, A.J. Understanding Kid's Digital Twin. In Proceedings of the 17th International Conference on Information and Knowledge Engineering (IKE), Las Vegas, NV, USA, 30 July–2 August 2018; CSREA Press: Las Vegas, NV, USA, 2018; pp. 41–46.
13. Verdouw, C.; Kruize, J. Digital Twins in Farm Management: Illustrations from the FIWARE Accelerators SmartAgriFood and Fractals. In Proceedings of 7th Asian-Australasian Conference on Precision Agriculture Digital, Hamilton, New Zealand, 16–18 October 2017; Procesion Agriculture Association New Zealand: Hamilton, New Zealand, 2017; pp. 1–5. [CrossRef]
14. Lu, Y.; Liu, C.; Wang, K.I.K.; Huang, H.; Xu, X. Digital Twin-driven smart manufacturing: Connotation, reference model, applications and research issues. *Robot. Comput. Manuf.* **2020**, *61*. [CrossRef]
15. Melesse, T.Y.; Di Pasquale, V.; Riemma, S. Digital Twin Models in Industrial Operations: A Systematic Literature Review. *Procedia Manuf.* **2020**, *42*, 267–272. [CrossRef]
16. Roy, R.B.; Mishra, D.; Pal, S.K.; Chakravarty, T.; Panda, S.; Chandra, M.G.; Pal, A.; Misra, P.; Chakravarty, D.; Misra, S. Digital twin: Current scenario and a case study on a manufacturing process. *Int. J. Adv. Manuf. Technol.* **2020**, *107*, 3691–3714. [CrossRef]
17. Schleich, B.; Anwer, N.; Mathieu, L.; Wartzack, S. Shaping the digital twin for design and production engineering. *CIRP Ann.* **2017**, *66*, 141–144. [CrossRef]
18. He, B.; Bai, K.-J. Digital twin-based sustainable intelligent manufacturing: A review. *Adv. Manuf.* **2020**, 1–21. [CrossRef]
19. Vachálek, J.; Bartalský, L.; Rovný, O.; Šišmišová, D.; Morháč, M.; Lokšík, M. The Digital Twin of an Industrial Production Line within the Industry 4.0 Concept. In Proceedings of the 21st International Conference on Process Control (PC), Štrbské Pleso, Slovakia, 6–9 June 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 258–262.
20. Boje, C.; Guerriero, A.; Kubicki, S.; Rezgui, Y. Towards a semantic Construction Digital Twin: Directions for future research. *Autom. Constr.* **2020**, *114*. [CrossRef]
21. Wang, B. The Future of Manufacturing: A New Perspective. *Engineering* **2018**, *4*, 722–728. [CrossRef]
22. Scott-Emuakpor, O.; George, T.; Beck, J.; Schwartz, J.; Holycross, C.; Shen, M.H.H.; Slater, J. Material Property Determination of Vibration Fatigued DMLS and Cold-Rolled Nickel Alloys. In Proceedings of the ASME Turbo Expo 2014: Turbine Technical Conference and Exposition, Düsseldorf, Germany, 16–20 June 2014; ASME: New York, NY, USA, 2014; Volume 7A.
23. DebRoy, T.; Zhang, W.; Turner, J.; Babu, S.S. Building digital twins of 3D printing machines. *Scr. Mater.* **2017**, *135*, 119–124. [CrossRef]
24. Majumdar, P.K.; FaisalHaider, M.; Reifsnider, K. Multi-Physics Response of Structural Composites and Framework for Modeling Using Material Geometry. In Proceedings of the 54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Boston, MA, USA, 8–11 April 2013.
25. Ricks, T.M.; Lacy, T.; Pineda, E.J.; Bednarcyk, B.A.; Arnold, S.M. Computationally efficient solution of the high-fidelity generalized method of cells micromechanics relations. In Proceedings of the American Society for Composites 30th Annual Technical Conference, East Lansing, MI, USA, 28–30 September 2015.
26. Qi, Q.; Tao, F.; Hu, T.; Anwer, N.; Liu, A.; Wei, Y.; Wang, L.; Nee, A.Y.C. Enabling technologies and tools for digital twin. *J. Manuf. Syst.* **2019**. [CrossRef]

27. Hoeber, H.; Alsem, D. Life-cycle information management using open-standard BIM. *Eng. Constr. Arch. Manag.* **2016**, *23*, 696–708. [CrossRef]
28. Chen, Y.; Jupp, J. Model-Based Systems Engineering and Through-Life Information Management in Complex Construction. Product Lifecycle Management to Support Industry 4.0. In Proceedings of the 15th IFIP International Conference on Product Lifecycle Management (PLM 2018), Turin, Italy, 2–4 July 2018; Springer: Cham, Switzerland, 2018; pp. 80–92.
29. Fernández Pérez, J.L.; Hernandez, C. *Practical Model-Based Systems Engineering*; Artech House: Boston, MA, USA, 2019.
30. Peters, S.; Fortin, C.; McSorley, G. A Novel Approach to Product Lifecycle Management and Engineering Using Behavioural Models for the Conceptual Design Phase. In Proceedings of the 16th IFIP International Conference on Product Lifecycle Management (PLM 2019), Moscow, Russia, 8–12 July 2019; Springer: Cham, Switzerland, 2019; pp. 159–169.
31. Mabkhot, M.M.; Al-Ahmari, A.M.; Salah, B.; Alkhalefah, H. Requirements of the smart factory system: A survey and perspective. *Machines* **2018**, *6*, 23. [CrossRef]
32. Nasir, M.F.M.; Hamzah, H.S. Supply chain management framework development for new multiple life cycle product development. In Proceedings of the 2016 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Bali, Indonesia, 5–7 December 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 812–816.
33. Lim, K.Y.H.; Zheng, P.; Chen, C.-H. A State-of-the-Art Survey of Digital Twin: Techniques, Engineering Product Lifecycle Management and Business Innovation Perspectives. *J. Intell. Manuf.* **2019**, *1–25*, 1313–1337. [CrossRef]
34. Awouda, A.; Aliev, K.; Chiabert, P.; Antonelli, D. Practical Implementation of Industry 4.0 Based on Open Access Tools and Technologies. In Proceedings of the 16th IFIP International Conference on Product Lifecycle Management (PLM 2019), Moscow, Russia, 8–12 July 2019; Springer: Cham, Switzerland, 2019; pp. 94–103.
35. Singh, V.; Willcox, K.E. Engineering Design with Digital Thread. *AIAA J.* **2018**, *56*, 4515–4528. [CrossRef]
36. Helu, M.; Hedberg, T.; Barnard Feeney, A. Reference architecture to integrate heterogeneous manufacturing systems for the digital thread. *CIRP J. Manuf. Sci. Technol.* **2017**, *19*, 191–195. [CrossRef]
37. Hedberg, T., Jr.; Feeney, A.B.; Helu, M.; Camelio, J.A. Toward a Lifecycle Information Framework and Technology in Manufacturing. *J. Comput. Inf. Sci. Eng.* **2017**, *17*. [CrossRef] [PubMed]
38. Mies, D.; Marsden, W.; Warde, S. Overview of Additive Manufacturing Informatics: "A Digital Thread". *Integr. Mater. Manuf. Innov.* **2016**, *5*, 114–142. [CrossRef]
39. Siedlak, D.J.L.; Pinon, O.J.; Schlais, P.R.; Schmidt, T.M.; Mavris, D.N. A digital thread approach to support manufacturing-influenced conceptual aircraft design. *Res. Eng. Des.* **2017**, *29*, 285–308. [CrossRef]
40. Bone, M.; Blackburn, M.; Kruse, B.; Dzielski, J.; Hagedorn, T.; Grosse, I. Toward an Interoperability and Integration Framework to Enable Digital Thread. *Systems* **2018**, *6*, 46. [CrossRef]
41. Phoenix Integration. Model Center MBSE. Available online: https://www.phoenix-int.com/product/mbse/ (accessed on 11 January 2021).
42. Finocchiaro, M. Demystifying Digital Thread and Digital Twin. 2017. Available online: https://www.linkedin.com/pulse/demystifying-digital-dilemmas-michael-finocchiaro (accessed on 7 January 2021).
43. Zheng, Y.; Yang, S.; Cheng, H. An application framework of digital twin and its case study. *J. Ambient. Intell. Humaniz. Comput.* **2018**, *10*, 1141–1153. [CrossRef]
44. Ivanov, D.; Dolgui, A. A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0. *Prod. Plan. Control* **2020**, 1–14. [CrossRef]
45. Tao, F.; Qi, Q.; Liu, A.; Kusiak, A. Data-driven smart manufacturing. *J. Manuf. Syst.* **2018**, *48*, 157–169. [CrossRef]
46. Borky, J.M.; Bradley, T.H. (Eds.) Protecting Information with Cybersecurity. In *Effective Model-Based Systems Engineering*; Springer International Publishing: Cham, Switzerland, 2019; pp. 345–404. [CrossRef]
47. Calder, A.; Watkins, S.G. *Information Security Risk Management for ISO27001/ISO27002*, 2nd ed.; IT Governance Publishing: Cambridgeshire, UK, 2010.
48. Katsikogiannis, G.; Mitropoulos, S.; Douligeris, C. An Identity and Access Management Approach for SOA. In Proceedings of the 2016 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Ajman, UAE, 12–14 December 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 126–131.
49. Detlef, G.; Gert, R.; Alexander, K.; Richard, L. Information Management in Product Development Workflows—A Novel Approach on the basis of Pseudonymization of Product Information. *Procedia CIRP* **2014**, *21*, 467–472. [CrossRef]
50. Curran, A. How Computer Design Software Delayed The Airbus A380. Available online: https://simpleflying.com/airbus-a380-computer-design-delay/ (accessed on 14 January 2021).
51. Kingsley-Jones, M. Farnborough First News: The Race to Rewire the Airbus A380. Available online: https://www.flightglobal.com/farnborough-first-news-the-race-to-rewire-the-airbus-a380/68529.article (accessed on 14 January 2021).
52. Fourgeau, E.; Gomez, E.; Hagege, M. Managing the Embedded Systems Development Process with Product Life Cycle Management. In *Complex Systems Design & Management Asia. Advances in Intelligent Systems and Computing*; Springer: Cham, Switzerland, 2016; Volume 426, pp. 147–158.
53. Ashtari Talkhestani, B.; Jung, T.; Lindemann, B.; Sahlab, N.; Jazdi, N.; Schloegl, W.; Weyrich, M. An architecture of an Intelligent Digital Twin in a Cyber-Physical Production System. *Automatisierungstechnik* **2019**, *67*, 762–782. [CrossRef]
54. Tao, F.; Cheng, J.; Qi, Q.; Zhang, M.; Zhang, H.; Sui, F. Digital twin-driven product design, manufacturing and service with big data. *Int. J. Adv. Manuf. Technol.* **2018**, *94*, 3563–3576. [CrossRef]

55. Zimmerman, P.; Gilbert, T.; Salvatore, F. Digital engineering transformation across the Department of Defense. *J. Def. Model. Simul.* **2017**, *16*, 325–338. [CrossRef]
56. Taylor, N.; Human, C.; Kruger, K.; Bekker, A.; Basson, A.; Taylor, N.; Human, C.; Kruger, K.; Bekker, A.; Basson, A. Comparison of Digital Twin Development in Manufacturing and Maritime Domains. In *International Workshop on Service Orientation in Holonic and Multi-Agent Manufacturing SOHOMA 2019*; Borangiu, T., Trentesaux, D., Leitão, P., Giret Boggino, A., Botti, V., Eds.; Springer: Cham, Switzerland, 2019; pp. 158–170. [CrossRef]
57. Ranchal, R.; Bhargava, B. Protecting PLM Data Throughout Their Lifecycle. Quality, Reliability, Security and Robustness in Heterogeneous Networks. In Proceedings of the 9th International Conference (QShine 2013), Greader Noida, India, 11–12 January 2013; Springer: Berlin/Heidelberg, Germany, 2013; pp. 633–642.
58. Ameri, F.; Dutta, D. Product Lifecycle Management: Closing the Knowledge Loops. *Comput. Des. Appl.* **2005**, *2*, 577–590. [CrossRef]
59. Vaz, C.R.; Selig, P.M.; Viegas, C.V. A proposal of intellectual capital maturity model (ICMM) evaluation. *J. Intellect. Cap.* **2019**, *20*, 208–234. [CrossRef]
60. Biahmou, A.; Stjepandić, J. Towards agile enterprise rights management in engineering collaboration. *Int. J. Agil. Syst. Manag.* **2016**, *9*, 302–325. [CrossRef]
61. Mason, A. Protection of Intellectual Property of the Plant Continuity through IT/OT Cyber Security Measures and Governance into Industrial Automation & Control Systems. Master's Thesis, The George Washington University, Washington, DC, USA, 2018.
62. Recamán Rivas, Á. Navantia's Shipyard 4.0 model overview. *Cienc. Tecnol. Buques* **2018**, *11*. [CrossRef]
63. Fraga-Lamas, P.; Fernandez-Carames, T.M.; Blanco-Novoa, O.; Vilar-Montesinos, M.A. A Review on Industrial Augmented Reality Systems for the Industry 4.0 Shipyard. *IEEE Access* **2018**, *6*, 13358–13375. [CrossRef]
64. Fernandez-Carames, T.M.; Fraga-Lamas, P.; Suarez-Albela, M.; Vilar-Montesinos, M. A Fog Computing and Cloudlet Based Augmented Reality System for the Industry 4.0 Shipyard. *Sensors* **2018**, *18*, 1798. [CrossRef]
65. Ramirez-Pena, M.; Abad Fraga, F.J.; Sanchez Sotano, A.J.; Batista, M. Shipbuilding 4.0 Index Approaching Supply Chain. *Materials* **2019**, *12*, 4129. [CrossRef]
66. Arrichiello, V.; Gualeni, P. Systems engineering and digital twin: A vision for the future of cruise ships design, production and operations. *Int. J. Interact. Des. Manuf.* **2019**, *14*, 115–122. [CrossRef]
67. Stanić, V.; Hadjina, M.; Fafandjel, N.; Matulja, T. Toward Shipbuilding 4.0—An Industry 4.0 Changing the Face of the Shipbuilding Industry. *Brodogradnja* **2018**, *69*, 111–128. [CrossRef]
68. Čelar, D. Augmented Reality for Naval Applications. *Nav. Eng. J.* **2017**, *129*, 55–57.
69. Morais, D.; Goulanian, G.; Danese, N. The Future Reality of the Digital Twin as a Cross-Enterprise Marine Asset. In Proceedings of the 19th International Conference on Computer Applications in Shipbuilding 2019, Rotterdam, The Netherlands, 24–26 September 2019; The Royal Institution of Naval Architects: London, UK, 2019; pp. 24–26.
70. Kuper, S. Top 5 for 2018: Defence Connect's Best SEA 1000 Stories. Available online: https://www.defenceconnect.com.au/maritime-antisub/3359-top-5-for-2018-defence-connect-s-best-sea-1000-stories (accessed on 29 June 2019).
71. 3DEXPERIENCE. *Industry Services Transition Factory: A Smooth Transition to the 3DEXPERIENCE*; Dassault Systèmes: Waltham, MA, USA, 2016. Available online: https://www.3ds.com/fileadmin/Products/Services/pdfs/services-Transition-Factory-flyer.pdf (accessed on 10 January 2021).
72. Hearn, M.; Rix, S. Cybersecurity Considerations for Digital Twin Implementations. *IIC J. Innov.* **2019**, 107–113. Available online: https://www.iiconsortium.org/news/journal-of-innovation-2019-nov.htm (accessed on 12 January 2021).

applied sciences

MDPI

Article

A Virtual Prototype for Fast Design and Visualization of Gerotor Pumps

Juan Pareja-Corcho [1,2], Aitor Moreno [2], Bruno Simoes [2], Asier Pedrera-Busselo [3], Ekain San-Jose [3], Oscar Ruiz-Salguero [1] and Jorge Posada [2,*]

[1] Laboratory of CAD CAM CAE, Universidad EAFIT, Cra 49 no 7-sur-50, 050022 Medellín, Colombia; jpareja1@eafit.edu.co (J.P.-C.); oruiz@eafit.edu.co (O.R.-S.)
[2] Vicomtech Foundation, Basque Research and Technology Alliance (BRTA), Mikeletegi 57, 20009 Donostia-San Sebastian, Spain; amoreno@vicomtech.org (A.M.); bsimoes@vicomtech.org (B.S.)
[3] Egile Innovative Solutions, Kurutz-Gain Polígono Industrial Pol., 12, 20850 Gipuzkoa, Spain; asier.pedrera@egile.es (A.P.-B.); ekain.sanjose@egile.es (E.S.-J.)
* Correspondence: jposada@vicomtech.org; Tel.: +34-943-309-230

Abstract: In the context of generation of lubrication flows, gear pumps are widely used, with gerotor-type pumps being specially popular, given their low cost, high compactness, and reliability. The design process of gerotor pumps requires the simulation of the fluid dynamics phenomena that characterize the fluid displacement by the pump. Designers and researchers mainly rely on these methods: (i) computational fluid dynamics (CFD) and (ii) lumped parameter models. CFD methods are accurate in predicting the behavior of the pump, at the expense of large computing resources and time. On the other hand, Lumped Parameter models are fast and they do not require CFD software, at the expense of diminished accuracy. Usually, Lumped Parameter fluid simulation is mounted on specialized black-box visual programming platforms. The resulting pressures and flow rates are then fed to the design software. In response to the current status, this manuscript reports a virtual prototype to be used in the context of a Digital Twin tool. Our approach: (1) integrates pump design, fast approximate simulation, and result visualization processes, (2) does not require an external numerical solver platforms for the approximate model, (3) allows for the fast simulation of gerotor performance using sensor data to feed the simulation model, and (4) compares simulated data vs. imported gerotor operational data. Our results show good agreement between our prediction and CFD-based simulations of the actual pump. Future work is required in predicting rotor micro-movements and cavitation effects, as well as further integration of the physical pump with the software tool.

Keywords: digital-twin; gerotor pump; hydraulic-systems; simulation; computer-aided design

check for updates

Citation: Pareja-Corcho, J.; Moreno, A.; Simoes, B.; Pedrera-Busselo, A.; San-Jose, E.; Ruiz-Salguero, O.; Posada, J. A Virtual Prototype for Fast Design and Visualization of Gerotor Pumps. *Appl. Sci.* **2021**, *11*, 1190. http://doi.org/10.3390/app11031190

Academic Editor: Andrew Y. C. Nee
Received: 16 December 2020
Accepted: 25 January 2021
Published: 28 January 2021

1. Introduction

Gerotor pumps play an important role in the aerospace industry, particularly in the processes of cooling, lubrication, and fuel boost and transfer. In other sectors, the gerotor pumps are operated in a wide range of applications, such as dosing and filling technologies in pharmacy and medicine, dispensing technologies and coating applications in manufacturing, among others. The popularity of such pumps in industrial applications arises from the fact that gerotor pumps represent a reasonable compromise in terms of compactness, reliability, cost, and versatility [1]. The working principle of a gerotor pump is based on the interaction between a pair of toothed gears with trochoidal envelope profiles. The relative movement between the profiles generates a series of chambers with varying volume that perform a cycle of suction and delivery actions (in interaction with input and output ports), thus effectively producing a volumetric flow (see Figure 1).

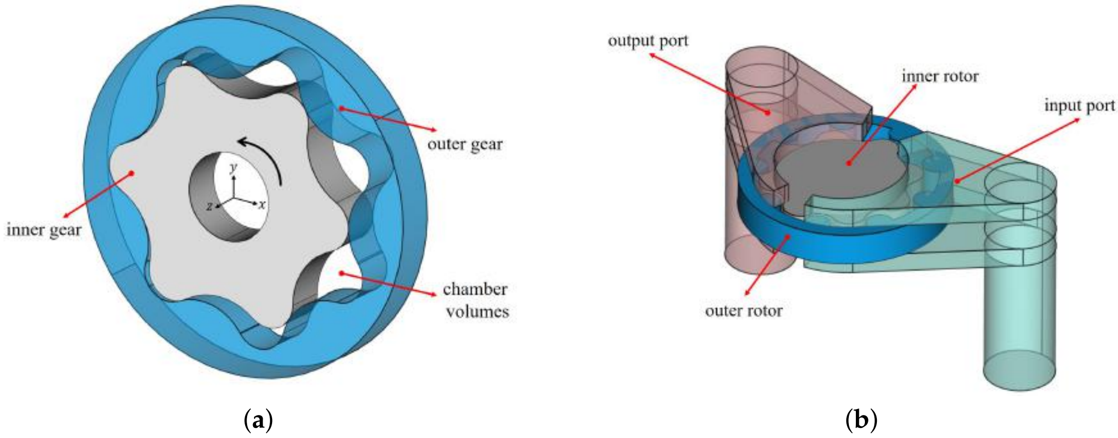


Figure 1. Gerotor pump general architecture: (**a**) inner and outer gear, (**b**) inlet/outlet disposition in pump.

The current design process for gerotor pumps commonly involves: (i) a *geometric modeling* step in a CAD environment, (ii) a *design verification* phase using fluid mechanics simulations to validate the efficiency and other desired characteristics of the pump, and (iii) a *physical testing* phase to verify the predicted characteristics of the pump in a real test bench once the design has been validated through a simulation tool. This process can be considerably time-consuming, due to the large amount of time that is required in the design verification stage. The design engineer must mesh the complex geometry of the volume chambers each time that the design is changed and perform a time-consuming simulation. In most design cases, the simulation of a geometric configuration takes up to a day to generate results. The described workflow hinders the effectiveness of rapid design methodologies or the easy testing of a large number of geometric configurations of the pump in a reasonable time.

Implementation. In this manuscript, we present the implementation of a virtual prototype of a Gerotor pump designed to be integrated with data measured in an experimental setup in order to improve the established design process. Our implementation does not constitute a full Digital Twin, but rather will be a step towards a fully functional Digital Twin tool that reproduces the behavior of the real pump. This virtual prototype allows for a rough design condition vs. performance appraisal, thus enabling the design and testing scenarios. Once the designer is satisfied with this approximated design vs. performance ratio, a more precise CFD simulation process would take place. An important current feature of the virtual prototype tool presented is the import and display of the sub-sequent CFD simulation results and experimental data measured in a real pump, for the benefit of the designer manufacturer and client. This feedback of the CFD simulation results might be included in a numerically oriented closed loop at the design stage. At the present time, we only report visual CFD data feedback. The implemented tool is able to use data that were measured in the experimental setup to feed the fast virtual prototype. Differences between the virtual prototype state variables and measured state variables allow for several activities: (a) to modify the pump design, (b) to control the actual pump, and (c) to feed satisfactory virtual prototype parameters into parametric or constraint-driven CAD models to obtain a full Boundary Representation of the Gerotor pump. Notice that (c) streamlines the design-for-gerotor process and avoids the need for a external CAD application.

The manuscript is divided, as follows: the Section 2 reviews the available literature in the context of physical simulation of gerotor pumps and Digital Twin implementations. Section 3 introduces the experimental setup of the pump, the lumped parameter model, and the virtual prototype tool. Section 4 presents the comparison between our predictions and a Computational Fluid Dynamics simulation used as ground truth. We do not address the comparison with respect to the experimental data, because we cannot measure the comparison variable in our experimental setup. Section 5 concludes the manuscript and discusses possible future developments in both the virtual prototype and its integration within a full Digital Twin tool.

2. Previous Works

In this section, we review the literature in two dimensions: (a) methods for fluid dynamics simulation in gerotor pumps and (b) implementations of virtual prototypes in Digital Twin oriented tools.

2.1. Fluid Mechanical Simulation

Several approaches have been proposed to simulate the performance of gerotor pumps, depending on the level of detail required. Most previous work relies on two methods for the fluid simulation: (i) lumped parameters models (LP) and (ii) computational fluid dynamics models (CFD), with each one exhibiting different performances regarding time and memory complexity.

CFD models: computational fluid dynamics models use specialized software to solve the Navier–Stokes equations in a discretized domain. CFD models can be classified in two categories: (i) two-dimensional (2D) models and (ii) three-dimensional (3D) models. Castilla et al. [2] and Houzeaux et al. [3] presented 2D CFD models for the simulation of rotary pumps that present accurate results with respect to an experimental setup. Recently, 3D simulations of the pump have been performed in order to analyze specific aspects of the pumps design, such as: (a) profile geometry optimization [4,5], (b) discharge coefficient calculation [6], and (c) fluid leakage due to clearances [7]. The main advantages of CFD based methods are: (i) detailed description of the fluid's behavior inside the cavity of the pump and (ii) very accurate prediction of the effect of cavitation and fluid–body interaction on performance. The main disadvantages of CFD methods are: (i) large simulation time and memory requirements, (ii) the requirement to remesh the entire domain in each step of the solution, and (iii) the difficulty to mesh appropriately the inter-teeth clearance domain [2].

LP models: lumped parameter models discretize the pump in a number of control volumes, where each CV (control volume) corresponds to a cavity of the pump. The mass and energy conservation equations are used to integrate the pressure in each control volume. The pressure inside each control volume will depend on the instantaneous volume of the chamber and net flowrate of fluid through its surface. Pellegri et al. [8,9] presented a simple lumped parameter model that was mounted on AMESIM software, coupled with a geometric module that calculates the instantaneous areas and volumes of the chambers. The results show good agreement between predicted and measured data. Shah et al. [10] presented a lumped parameter model in AMESIM software for the prediction of cavitation effects on the pump simulation; the results show that the model is accurate in predicting the effect of cavitation phenomena on the overall performance of the system. The main advantages of the lumped parameter approach are: (i) the low time and memory complexity and (ii) the flexibility to integrate with larger hydraulic circuits [1,8,9]. The main disadvantages of the lumped parameter approaches are: (i) the results are coarse with respect to CFD methods and (ii) calibration of the model vs. experimental data is needed, which makes this approach unsuitable for detailed analysis of local behavior of fluid [11].

2.2. Digital Twins and Virtual Prototypes in Gerotor Applications

Digital Twins are virtual abstractions of physical products, processes, or phenomena very commonly used in the context of Industry 4.0 [12]. Digital Twins are a valuable tool in digital design and manufacturing, as they allow for prediction of system performance and simulation/optimization. Relatively few applications of Digital Twin methodology are found in industrial contexts [13], opening opportunities for wider adoption of Digital Twins in industries, such as fluid power systems. The use of accelerated coarse simulations for fast decision making, although not being entirely similar to the concept of Digital Twin, is being explored in other industrial contexts, such as quality control in manufacturing [14,15].

The lumped-parameter models that have been cited in the previous sections are usually implemented in specialized commercial software. This restriction limitates their feasibility towards a fully functional Digital Twin tool that integrates data from an experimental test bench. In the case of lumped parameter models, the design engineer must express the pump in a CAD environment and then import the geometric data into a differential equation solver (e.g., AMESIM [9]). In the case of CFD models, several commercial codes are used in the solution of the Navier–Stokes equations, including PumpLinux, ANSYS Fluent, and CFX (all appearing in Ref. [11]). So far, we have found no standalone fully-integrated implementations of gerotor pump simulation environments that suits our design needs.

2.3. Conclusions of Literature Review

Two approaches are commonly used in the context of gerotor pump simulation: (i) Lumped Parameter (LP) models and (ii) Computational Fluid Dynamics (CFD) models. The lumped parameter models allow for the fast simulation of pump performance at the price of loss of accuracy and detail. CFD models allow for very accurate simulation of pump performance with detailed information regarding in-chamber heterogeneity, at the expense of large simulation time and complexity. Because of its accuracy, it is common to use CFD as a ground truth value for pump experiments when no experimentally obtained comparison data are available. We use CFD as our point of comparison for the reasons expressed above and the availability of CFD software simulation. Furthermore, we found that implementations of both approaches are: (i) dependant on proprietary commercial software and (ii) not easily integrated with other standalone non-commercial design and optimization tools in the context of Digital Twins tools.

As a response to such shortcomings, we present the implementation of a virtual prototype for a gerotor pump, which also allows for the integration of measured data, thus enabling the functioning of a Digital Twin. Our implementation: (1) integrates pump design, fast approximate simulation, and result visualization processes, (2) does not require an external numerical solver platforms for the approximate model (as other approaches do), (3) allows for the fast simulation of gerotor performance, and (4) feeds the simulation model with data measured in an experimental setup to improve the accuracy of the model. Several variables can be used to assess the pump behavior, including, among others, maximum pressure, torque, and power. We use the maximum pressure in the pump as our comparison variable since (1) it is a variable of interest for the pump manufacturer and the variable that our model predicts and (2) the comparison with other variables (e.g., torque) would require specialized proprietary software. This software is not available to the industry manufacturer. See Implementation in the Introduction section.

3. Methodology

For our Digital Twin (DT) implementation, we have defined a three step workflow: (i) the design engineer inputs the values for the parameterization of pump design into the Geometry Configurator, (ii) the geometric model of the pump (inner and outer gears) is generated, and (iii) the virtual prototype performs the fast simulation of the pump with the geometric data and experimental data being measured from the test bench (see Figure 2). Please see the Abbreviations section immediately before the References section to find the meaning for the symbols used.

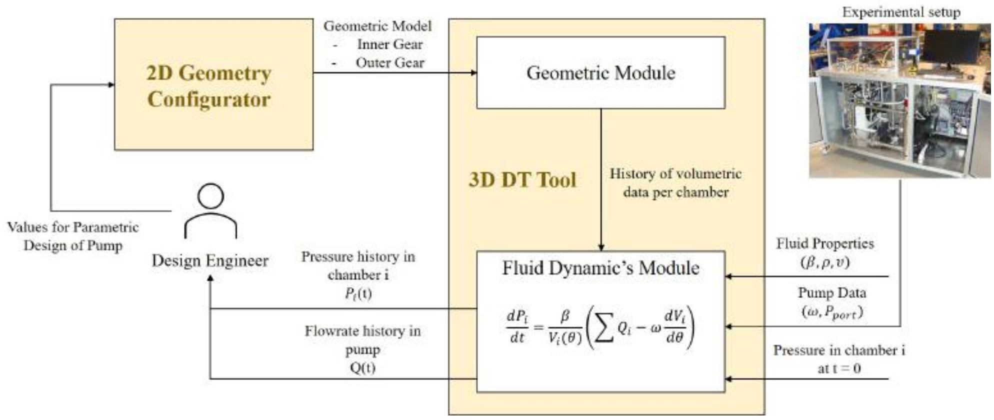
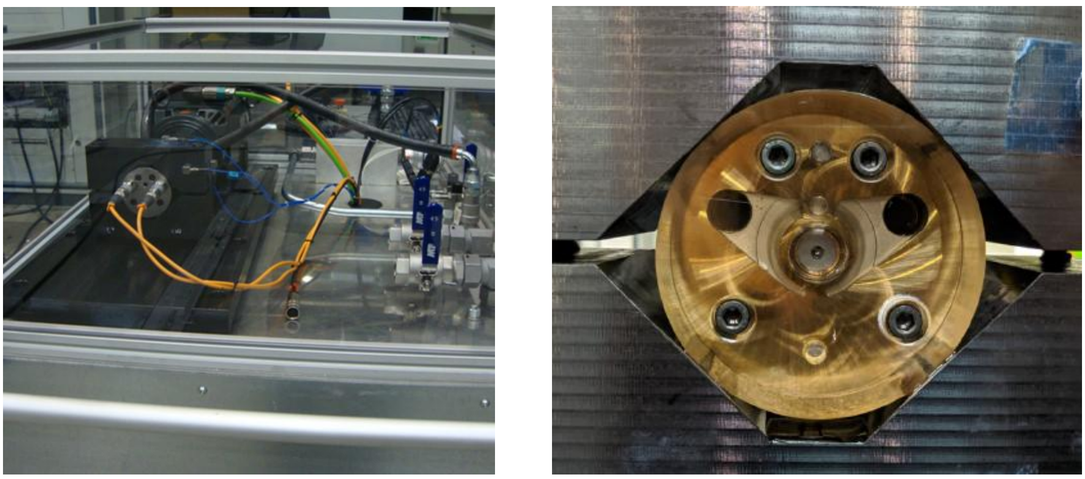


Figure 2. Implemented Tool Architecture and Workflow.

This section is divided in three parts: (1) we present the experimental setup where we show the data collection setup used to feed the virtual prototype tool, (2) we explain the geometric model, showing the generation of the gerotor geometric model and the calculation of geometric quantities such as the history of chamber volumes and areas, (3) we present the fluid dynamics module, where we lay out the foundations of the simulation model and (4) we discuss the software tool that integrates the virtual prototype model, including 3D visualization, with the data being collected from the experimental setup and external CFD simulations.

3.1. Experimental Setup

The contact point between the physical pump and the computational modeling is the testing bench. The experimental setup hosts the sensors that are used to collect performance data and feed it to the simulation model. Figure 3 presents the testing bench setup used and a manufactured pump mounted in the bench with a translucent cover.



(**a**) (**b**)

Figure 3. Experimental setup: (**a**) testing bench and (**b**) physical pump mounted on the test setup (translucent cover).

Pressure sensors that are located in the input and output ports measure the pressure value to be fed to the virtual prototype (as shown in Figure 2). Further integration of the data that were collected in the testing bench for variables other than pressure with the simulation model is still to be addressed.

3.2. Geometric Model

The internal profile of the gerotor is generated according to the parameterization proposed by Ref. [16]:

$$x_i(\alpha_{pc}) = R_2 \cos\left(\frac{1}{Z-1}\alpha_{pc}\right) \pm e \cos\left(\frac{Z}{Z-1}\alpha_{pc}\right) - \frac{S}{m}\left[R_2 \cos\left(\frac{1}{Z-1}\alpha_{pc}\right) \pm r_2 \cos\left(\frac{Z}{Z-1}\alpha_{pc}\right)\right] \quad (1)$$

$$y_i(\alpha_{pc}) = -R_2 \sin\left(\frac{1}{Z-1}\alpha_{pc}\right) \mp e \sin\left(\frac{Z}{Z-1}\alpha_{pc}\right) + \frac{S}{m}\left[R_2 \sin\left(\frac{1}{Z-1}\alpha_{pc}\right) \pm r_2 \sin\left(\frac{Z}{Z-1}\alpha_{pc}\right)\right] \quad (2)$$

$$m = \sqrt{r_2^2 + R_2^2 \pm 2 r_2 R_2 \cos \alpha_{pc}} \quad (3)$$

where the parameter $\alpha_{pc} \in [0, 2\pi]$ corresponds to the turning angle of the rotor. Figure 4 shows the resulting shape of the internal profile as the trace of contact point P', whose position is determined by the radii R_2 and r_2, the eccentricity e, and the number of chambers Z. In Figure 4, the external profile is determined by a set of outer circumferences that are truncated by a larger cutting circumference.

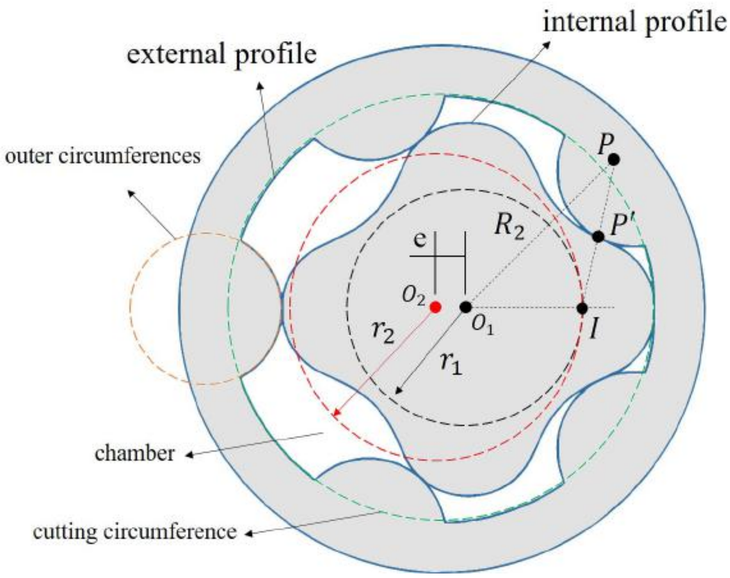


Figure 4. Parameterization of internal profile shape as in Equation (2).

This construction method of the external profile, even though simple and widely used, limits the performance of the pump, as the resulting shape does not mesh perfectly with the internal profile shape [17]. We have implemented an additional method to build the external profile as the conjugated curve of the internal shape.

Suppose the curve C that corresponds to the internal profile (Equations (1) and (2)) is put through a series of affine transformations that are defined by the rolling without slipping of the circumference defined by r_1 with respect to the circumference defined by r_2 in Figure 4. Subsequently, the external profile shape will be defined as the envelope curve of the locus of C as it moves through the rotation domain.

Figure 5 shows the locus of curve C, as generated by the movement of the circumferences. The envelope curve of the locus can be used as the external profile shape, with the advantage that by using this external shape both curves mesh perfectly, hence improving the performance by avoiding fluid recirculation. Once the inner and outer profile shapes are defined, the geometric quantities of each chamber are calculated by sampling both the internal and external curve to form a closed polygon, finding the area A_i and perimeter P_i of the polygon corresponding to chamber i.

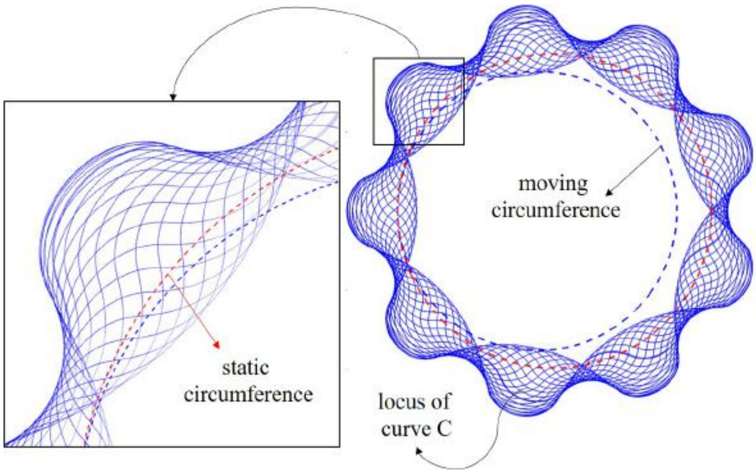


Figure 5. Locus of the internal profile curve *C*.

3.3. Fluid Dynamics Module

We discretize the flow domain in several control volumes to obtain a lumped parameter model of the pump, as shown in Figure 6. We assume the fluid properties within each control volume (CV) to be homogeneous, but not constant in time, effectively treating each control volume as the basic domain of simulation. Notice that, as the pump rotates, the geometry of the control volumes changes; therefore, the model requires a constant update of the geometric calculations for each control volume (area, perimeter) as the position of the pump changes (see Figure 2).

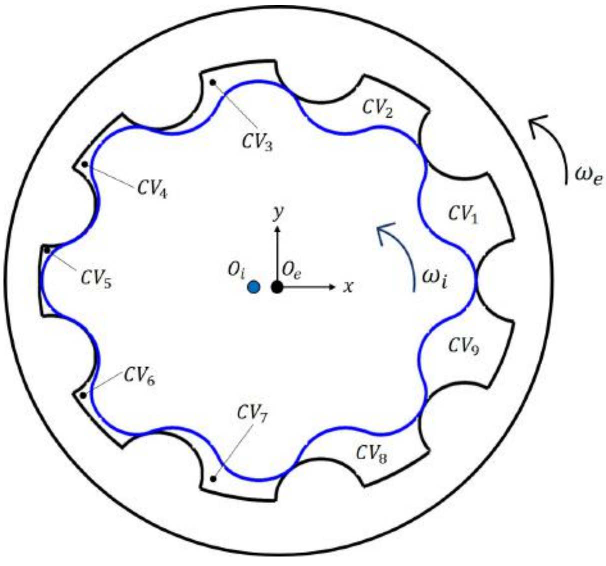


Figure 6. Control Volume discretization of the gerotor pump.

By the principle of the conservation of mass and energy, along with Reynold's transport theorem, it is possible to derive an expression for the change of pressure within a control volume:

$$\frac{dp}{dt} = \frac{\beta_{eff}(p)}{V(\theta)}\left(\sum Q_i - \omega\frac{dV}{d\theta}\right) \tag{4}$$

We omit the derivation of such an expression, since it is beyond the scope of our paper, the interested reader can find a thorough explanation in Ref. [11]. The net flowrate that flows through the boundary of a control volume needs to be calculated in order to integrate Equation (4), and since the volume and volume derivative at angle θ are provided by the geometric module. We consider two types of flows through the boundary of a control volume (Figure 7):

1. *Input/Output flow*: fluid flowing from the input port to the inside of the pump (charge) or from the inside of the pump to the output port (discharge).
2. *Fluid leak flow*: fluid flowing from one chamber to another due to imperfect sealing that results from manufacturing defects and design constraints.

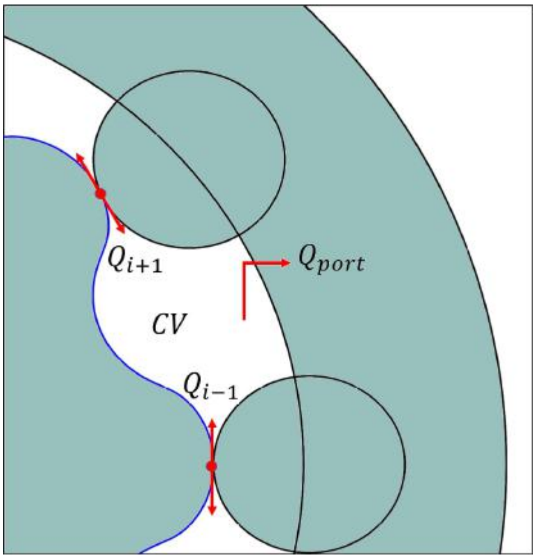


Figure 7. Types of flows through the boundary of a control volume.

The flow between the control volume and the input/output port is modeled as the flow through a variable geometry orifice subject to a difference in pressure. The pressure at the input and output ports is fed to the numerical model with data that were collected from the experimental testing bench, while the control volume pressure varies according to Equation (4). The flowrate in this situation can be obtained as:

$$Q_{in} = C_d A_{i,in} \sqrt{\frac{2(P_i - P_{in})}{\rho_{eff}}} \tag{5}$$

$$Q_{out} = C_d A_{i,out} \sqrt{\frac{2(P_i - P_{out})}{\rho_{eff}}} \tag{6}$$

Notice that, since a control volume only interacts with one of the ports at any given time, the net port flowrate Q for a control volume is equal to Q_{in} or Q_{out}, depending on the position of the control volume at the time of analysis. The calculation of the discharge coefficient C_d depends on the value of the Reynold's number Re and the hydraulic diameter D_h at such a time. The hydraulic diameter D_h and the Reynold's number Re are calculated, as follows:

$$D_h = \frac{4A_i(\theta)}{P_i(\theta)} \tag{7}$$

$$R_e = \frac{D_h}{\nu} \sqrt{\frac{2\Delta P}{\rho_{eff}}} \tag{8}$$

Finally, the discharge coefficient C_d is estimated while using an experimental expression (Ref. [11]):

$$C_d = C_{d\,\max} \tanh\left(\frac{2R_e}{R_{ecrit}}\right) \tag{9}$$

where $C_{d\,\max}$ is the maximum discharge coefficient and R_{ecrit} is the critical Reynolds number, which indicates the transition between laminar and turbulent regime. Values for constants $C_{d\,\max}$ and R_{ecrit} can be found in the literature as a function of conditions of the pump [8]. This flow between adjacent control volumes that should be nominally tight is enabled by the small gap between the rotors at their maximal approximation position. These gaps are necessary for ensuring rotation and limit friction and wear. The resulting fluid migration between adjacent chambers is caused by:

1. Difference of pressure between adjacent control volumes (Poiseuille flow).
2. Difference in angular speed between inner and outer rotor (Couette flow).

Typically, the gap between rotors at contact points is very small when compared to the overall size of the pump. The curvature radii at the throat are much larger than the throat gap. Therefore, (1) and (2), above, may be modeled by assuming that the approaching teeth form a constant clearance gap between two parallel plates (Figures 8 and 9).

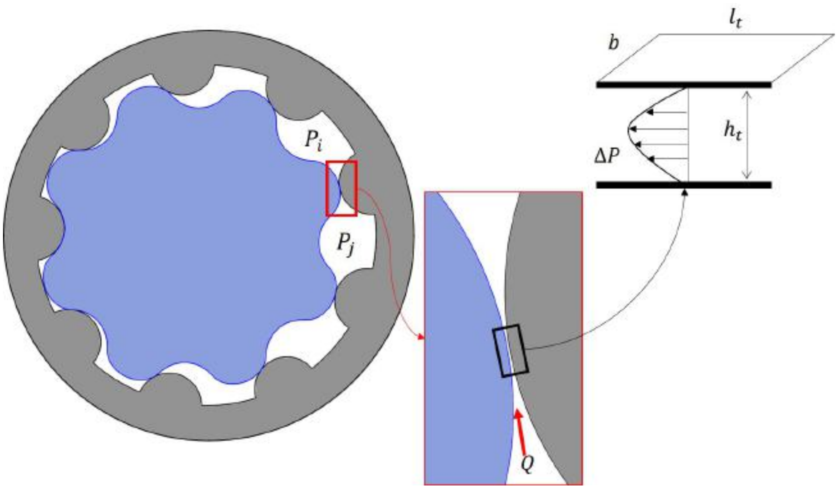


Figure 8. Pouiseuille flow between adjacent control volumes.

Figure 8 shows the working principle of the Poiseuille flow in the pump case, where two static plates of length l_t and width b are separated by a distance h_t. The difference in pressure between adjacent control volumes induces a flow Q_p that can be obtained as:

$$Q_p = b\frac{\Delta P(\frac{h_t}{2})^3}{12\mu L} \tag{10}$$

Notice that width b corresponds to the length of the pump profiles in the z direction. The distance h_t is estimated by the geometric module as it may vary for each contact point throughout the rotation of the rotors. As length l_t cannot be directly measured in the geometric model, we estimate l_t as a function of h_t. Starting from the point of minimum distance h_t, we move outwards through the profile curves to the point where the distance between rotors is $h_t* = (1+\epsilon)h_t$. Once such points are found, the length l_t is assumed to be the Euclidean distance between the points found. This approximation has shown to be effective for values of ϵ around 0.1 [8].

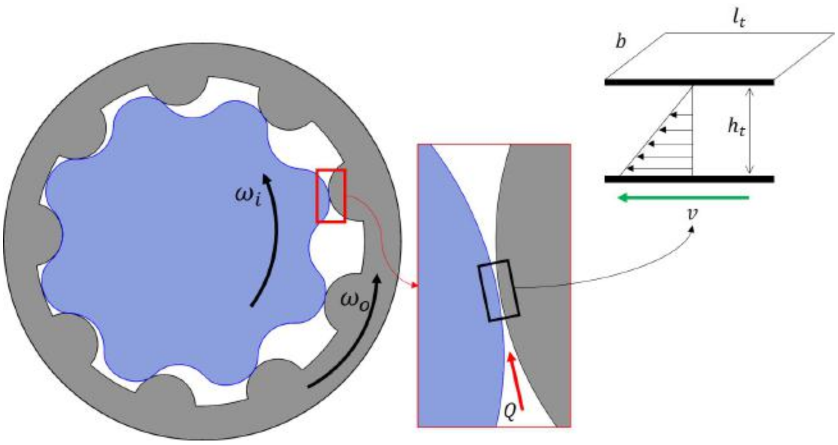


Figure 9. Couette flow between adjacent control volumes.

Figure 9 shows the working principle of the Couette flow. Two parallel plates having relative velocity with respect to each other produce a flowrate between them through the viscosity of the fluid and shear stress induced by the relative movement of the plates. The Couette and Poiseuille flows (Figures 8 and 9) both result in a fluid exchange between adjacent control volumes, therefore affecting the net flowrate through their borders and their pressures. Finally, as the pump usually operates in a low pressure range, the variance of effective fluid properties (bulk modulus, density) in hydraulic oil with respect to the instantaneous pressure [18] inside a control volume must be taken into account:

$$\beta_{\text{ef}} = \frac{\beta_{\text{oil}}}{1 + \alpha \cdot \left(\frac{p_0}{p}\right)^{\frac{1}{\kappa}} \cdot \left(\frac{\beta_{\text{oil}}}{\kappa \cdot p} - 1\right)} \tag{11}$$

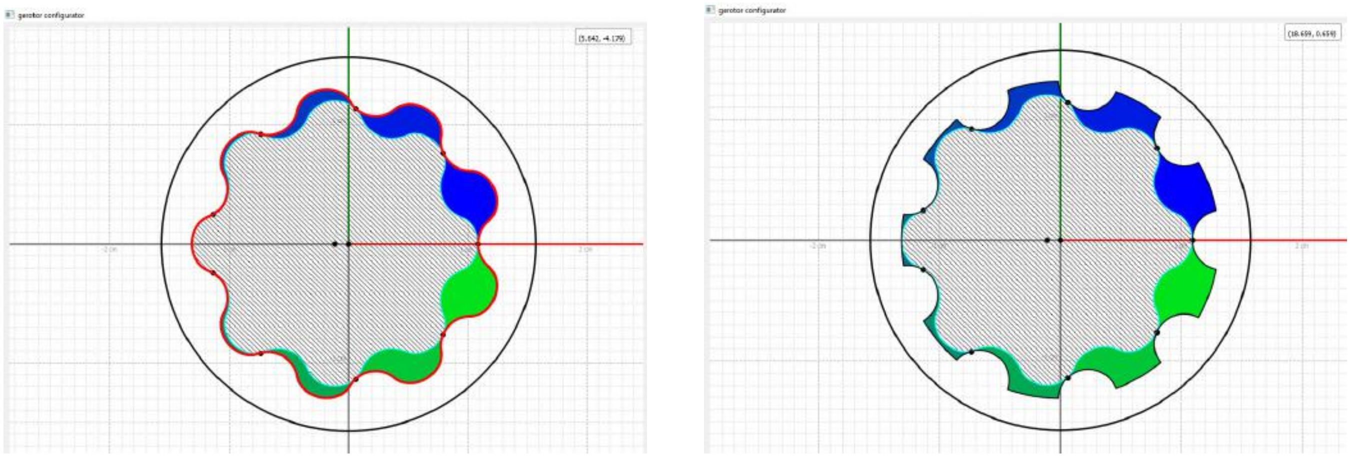
$$\rho_{\text{ef}} = \frac{\alpha \cdot \rho_{\text{air},0} + (1-\alpha) \cdot \rho_{\text{oil},0}}{\alpha \cdot \left(\frac{p_0}{p}\right)^{\frac{1}{K}} + (1-\alpha) \cdot \left(1 + \frac{m \cdot (p - p_0)}{\beta_{oil}}\right)^{-\frac{1}{m}}} \tag{12}$$

where β_{oil}, ρ_{air}, and ρ_{oil} are the properties of the oil and air at atmospheric conditions, respectively. p_0 is the atmospheric pressure, α is the void fraction, and κ is the polytropic constant of air.

3.4. Software Tool

The Digital Twin (DT) tool implements two functionalities: a *2D geometry configurator* and a *3D data tool*. The 2D geometry configurator allows for the design engineer to define a new geometry for the profiles of the inner and outer gears according to a set of parametric variables. The 3D data tool converts the model that is defined in the 2D geometry configurator to a full B-Rep model for simulation and visualization purposes of both (i) data simulated from our pre-CFD simulation model and (ii) data imported from CFD simulations or test bench.

Figure 10 shows the visualization of the parameterized pump in the interface of the 2D geometry configurator tool. The tool allows for the design engineer to input the desired set of values for the parametric variables of Equation (2). Our application automatically generates both the conjugated design (Figure 10a) and the classic design (Figure 10b). Both types of design, as well as other geometric configurations, are easily explorable in the 2D visualization canvas.



(**a**) (**b**)

Figure 10. Visualization of parameterized pump in the two-dimensional (2D) geometry configurator interface: (**a**) pump with conjugated external profile and (**b**) pump with classic external profile.

In addition to the discretized geometric model of the inner and outer gears, the geometry configurator calculates the shape of the resulting chambers for any given angle of the rotation of the pump. The configurator allows the user to interactively rotate the pump position with a slider, as well as to visualize and record the change in the area of each one of the chambers.

The ports and chambers geometries must be well-aligned for the correct calculation of the intersection area between the chambers and the ports. The geometry of the input/output ports (shown in Figure 11) of the pump is calculated from the geometry of the inner and outer gear geometries in order to ensure the alignment. The geometries calculated by this application (gear, chamber, and port geometry) are automatically imported into the 3D data tool for simulation purposes. The CAD models are also exported to external files.

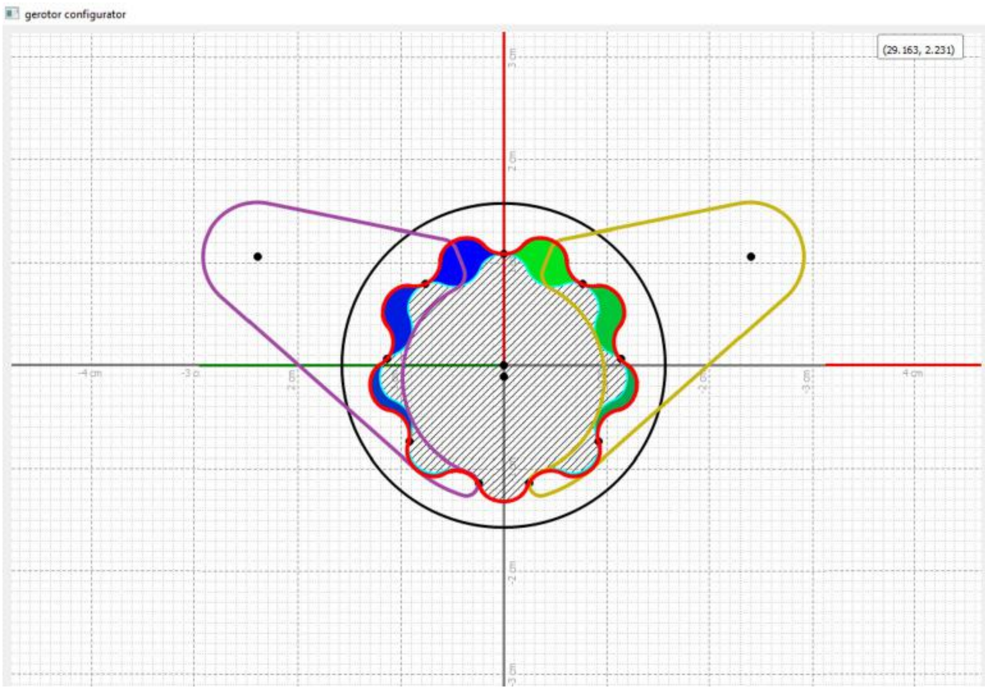


Figure 11. Gerotor design with input/output ports.

Figure 12 shows the interface that is devised for the toolkit that was introduced in [19,20] to enable the virtual prototype. The interface is built with Qt™ library and it comprises several panels: (i) the *3D visualization panel*, which allows for the visualization of the pump geometry (gears, ports, and chambers) and the animation showing the rotation of the pump through a pumping cycle, (ii) a *hierarchy panel* that shows the hierarchy tree of the geometric model and allows the user to select and highlight geometric entities, (iii) a *settings panel* in which the user inputs the fluid properties, operating conditions, and initial pressure conditions that are necessary to perform the fast pre-CFD simulation (see Figure 2), (iv) a *simulated data panel* that displays the results from the fast simulation model, and (v) a *measured data panel* that displays the results of the variables that are imported to the virtual prototype tool, whether measured in the test bench or simulated by the CFD software.

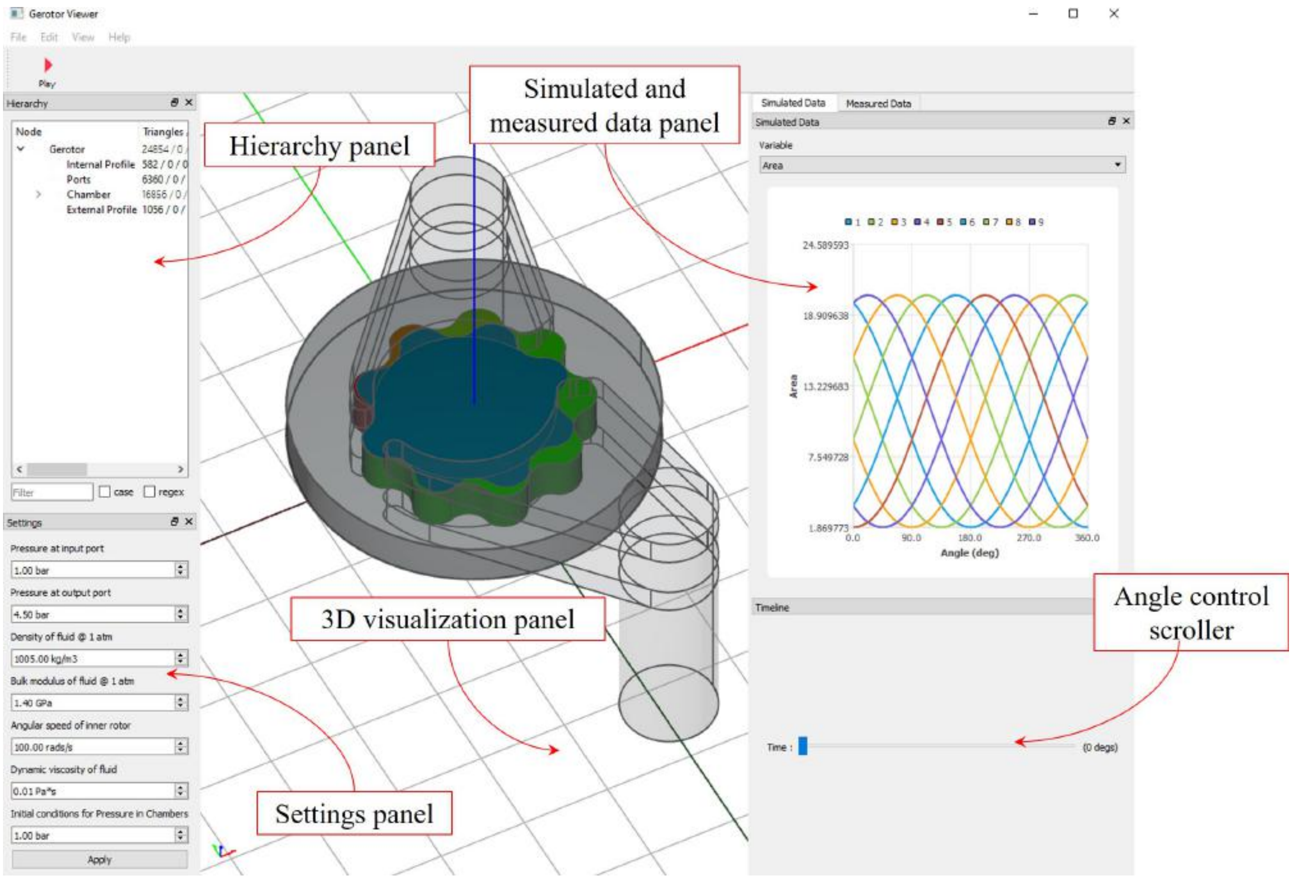
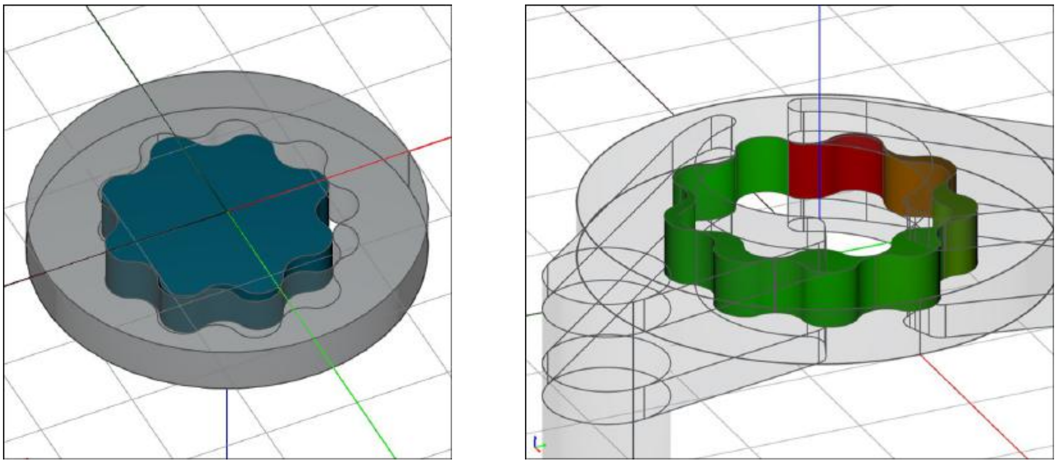


Figure 12. Three-dimensional (3D) data tool interface.

The 3D visualization panel enables design engineers to examine the three-dimensional behavior of the pumps for the parameters that are defined in the Geometry Configuration interface. Notice that the shape of the chambers must be recalculated each time the pump is set to a new angle position. Figure 13 shows the high degree of flexibility that can be achieved in the visualization of the different components of the pump geometry.

In Figure 13b, the color of each chamber is related to the pressure that is calculated by the simulation model. The tool also allows for the design engineer to import data from the CFD simulations and test bench tests with the objective of integrating and comparing performance data (both predicted and measured) in the same software environment. Figure 14 shows the data for a measured variable that was imported from an external CFD simulation, as visualized within the software tool. It is possible to import external data in the form of the widely used CSV format or while using a special text formatting that was specifically tuned for our application.



(**a**) (**b**)

Figure 13. Geometry visualization: (**a**) inner and outer gear of the pump, (**b**) fluid chambers with transparent ports.

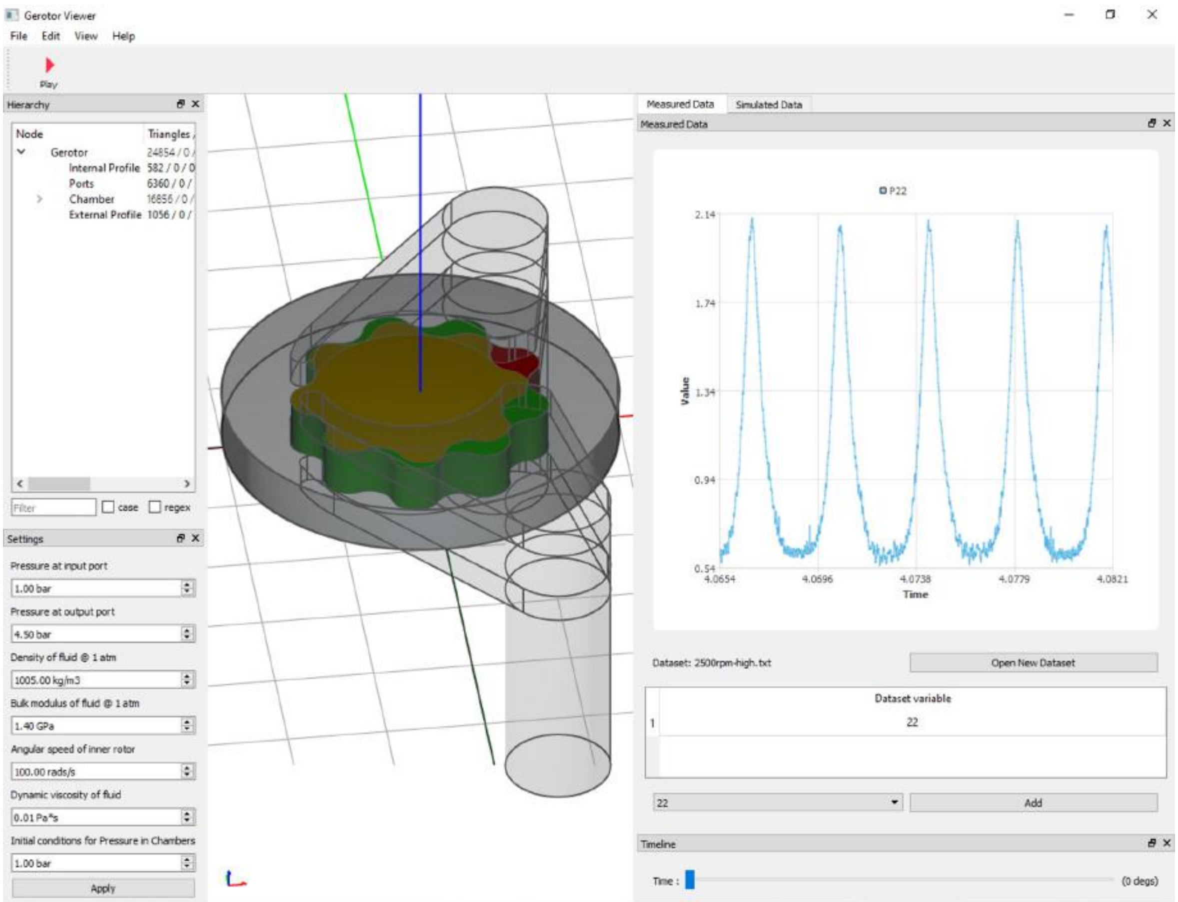


Figure 14. Example of the pressure data collected from the experimental setup sensors.

4. Results

We present the results of a design case while using the presented virtual prototype tool to design and simulate a gerotor pump. We also compare the obtained results with the data imported from an external CFD simulation for the same geometry, fluid properties, and operating conditions. Figures 15 and 16 present the volumetric data that were calculated by our tool for the pump design used in this test run.

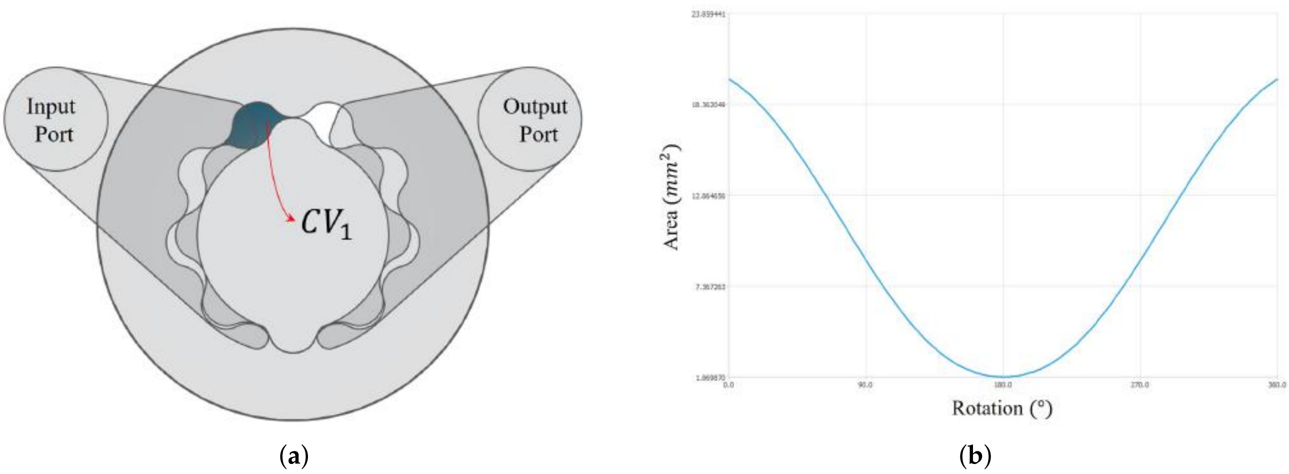


Figure 15. Volumetric data results from the virtual prototype: (**a**) profile of geometry with highlighted chamber (CV_1) and (**b**) history of area in a z-cut for selected chamber.

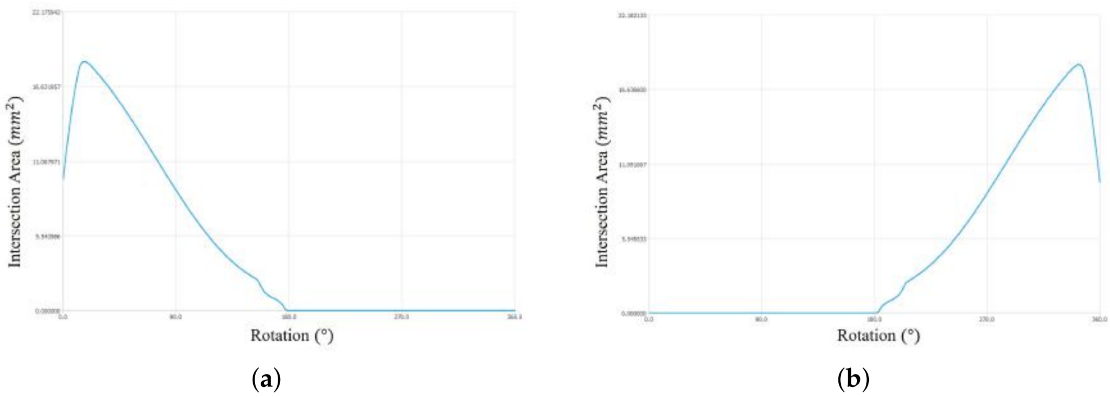


Figure 16. Volumetric data results from the virtual prototype: (**a**) history of intersection area between chamber CV_1 and input port and (**b**) history of intersection area between chamber CV_1 and output port.

Figure 15a shows the selected analysis chamber. In this case, we have selected the maximum volume chamber when the pump is at initial position $t = 0$, even though our tool can perform the geometric analysis of all chambers simultaneously. Figure 15b shows the area evolution for the analysis chamber. Notice that it starts with the maximum value and diminishes until it reaches the minimum value for chamber area around halfway through a revolution. The minimum value for the area is not exactly zero because of the gaps in the meshing between the internal and external rotors described in the previous section. The area then rises until it reaches the maximum value once the revolution of the pump is completed. Figure 16a shows the intersection area between the analysis chamber and the input port (area through which the working fluid enters the pump). Notice that the intersection area increases at first, because, at initial position, the port and the chamber are only partially overlapped (as seen in Figure 15a). When the chamber is no longer overlapping with the input port and has started discharging fluid (overlapping with the output port), the intersection area with the input port becomes zero. The same analysis corresponds to the intersection area between the chamber and output port shown in Figure 16b. Figure 17 shows the history of areas for all chambers in the pump; the periodicity is explained by the cyclic design of the pump.

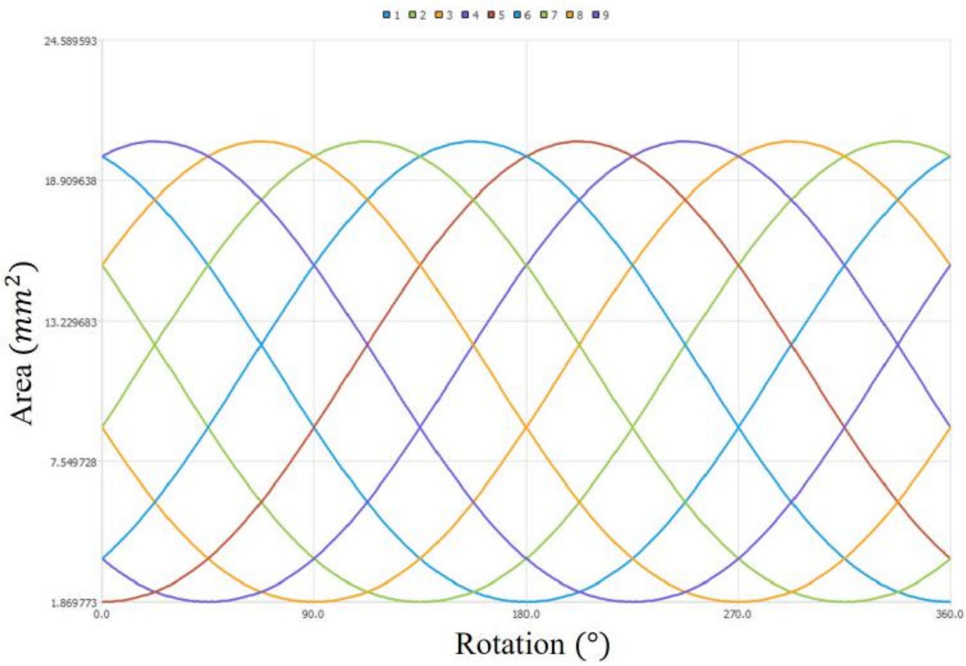


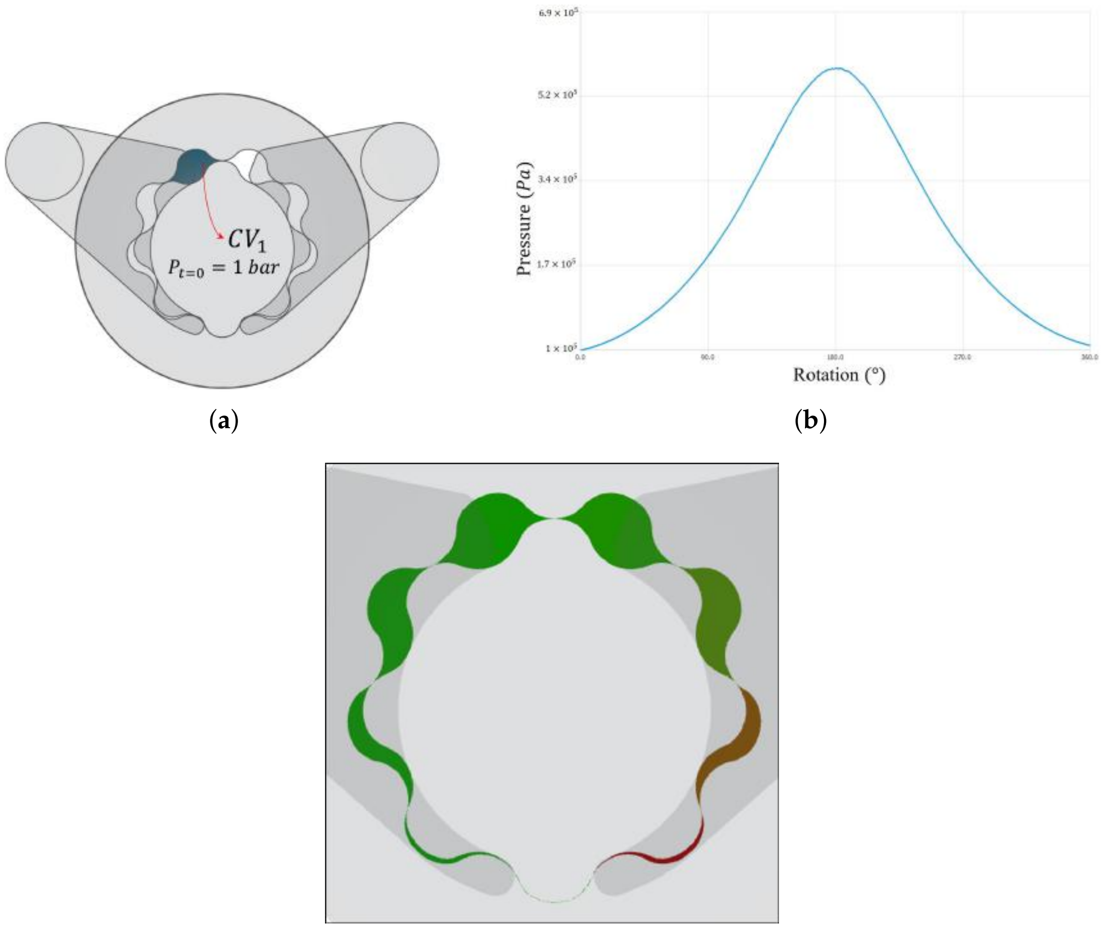
Figure 17. History of area for all 9 chambers in the pump.

We now discuss the results of a fluid dynamics simulation with our virtual prototype tool. The fluid properties, operating conditions and initial pressure conditions that are used for the simulation presented, are described in Table 1:

Table 1. Simulation conditions for test operating point.

Variable	Value	Units
Pressure at input port (p_{in})	1	bar
Pressure at output port (p_{out})	4.5	bar
Density of fluid at 1 atm (ρ_{oil})	1005	kg/m^3
Bulk modulus of fluid at 1 atm (β_{oil})	1.4	GPa
Angular speed of inner rotor (ω)	100	rad/s
Dynamic viscosity (μ)	0.01	Pa*s
Pressure in chambers at $t = 0$	1	bar

Figure 18 shows the history of calculated pressure inside an analysis chamber. Notice that the analysis chamber CV_1 (Figure 18a) is initially near the maximum volume position and, therefore, the initial pressure $P_{t=0}$ will be low (near input port pressure). As the pump rotates, the volume of the chamber reduces and increases the pressure inside the chamber. The maximum value of the pressure in chamber is reached in the minimum volume position and, once the analysis chamber enters the discharge cycle, the pressure starts to reduce as a result of the discharge of fluid and increase in volume of the chamber itself.



(**c**)

Figure 18. Pressure results from virtual prototype: (**a**) profile of geometry with highlighted chamber (CV_1), (**b**) history of pressure in chamber CV_1, and (**c**) pressure distribution in pump after a full revolution, as seen in Digital Twin (DT).

Figure 18c shows the pressure distribution in pump after a full revolution. Notice that there is a peak in pressure in the chambers near the minimum volume position and the pressure starts decreasing after the chamber intersects with the output port and goes through the discharge cycle. The discretization that was used in our virtual prototype allows for the tracking of the history of pressure for each individual chamber. However, the discretization used in the CFD methods make it impossible to track the history of pressure for an individual chamber. This is because, in CFD, the entire discretization (mesh) is updated in every time step (remeshing) to meet convergence requirements [21]. Therefore, for the purposes of comparison against a benchmark (i.e., CFD data), we use the maximal pressure across the entire fluid domain (all chambers). The CFD simulations are currently used in most of the design processes as an accurate prediction of the pump's performance. Therefore, they are a valid point of comparison for our implementation. Figure 19 shows the results of maximum pressure in pump, as predicted by both CFD simulation and by our implementation.

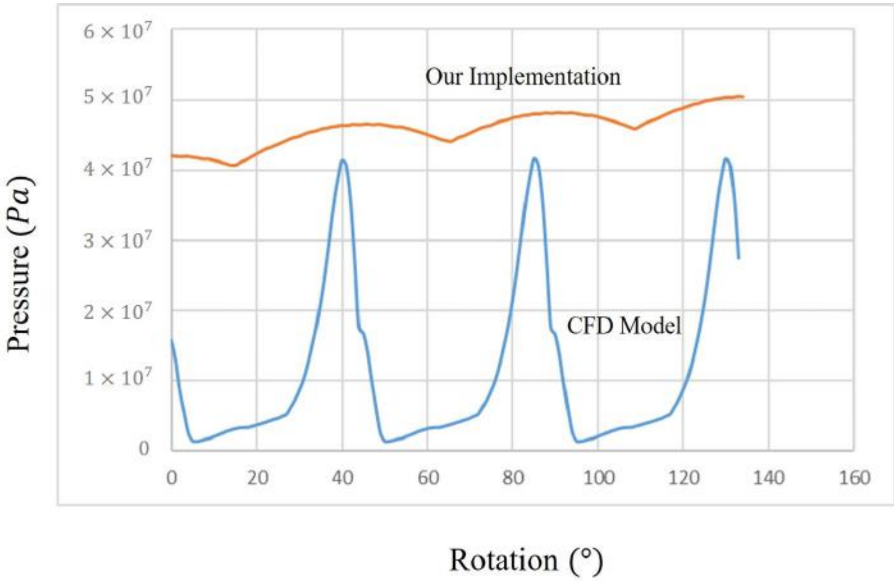


Figure 19. Virtual Prototype vs. computational fluid dynamics models (CFD) maximum predicted pressure in gerotor pump.

Our implemented virtual prototype is able to estimate the maximum pressure within the pump with a relative error of 21%. Our virtual prototype fails to reproduce the amplitude of the pressure oscillation in the pump. We believe that this shortcoming is due to the used assumption of homogeneous pressure in each one of the control volumes. The CFD model does not need such an assumption, and it is able to take into account variations of pressure within the chambers. Still, the maximum pressure in the pump as it rotates through an entire revolution is an important indication of the pump's performance, and it is reasonably estimated by our virtual prototype.

Table 2 shows a comparison of pre-processing and simulation times for both a CFD simulation and a simulation with our virtual prototype. The pre-processing time of the CFD simulation includes the generation of CAD models and mesh generation. The pre-processing time in our virtual prototype includes the parameter input for the geometry configuration tool and the automatic generation of B-Rep models. The main advantage of our virtual prototype is the much lower processing and simulation time, while still providing valuable performance information to the design engineer. The economic impact of our prototype upon the maker's functioning is not available at the present time, because of the fact that the design process, at the maker's facilities, is intertwined with other products and processes.

Table 2. Comparison of pre-processing and simulation times for the CFD simulation and our virtual prototype.

Task	CFD Time	Our Implementation Time
Pre-Processing of Geometry	1 h	<5 min
Simulation	9 h	<5 min

5. Conclusions and Future Work

In this manuscript, we have presented the implementation of a virtual prototype tool in the context of gerotor pump design, a component that has been widely used in different industries and that usually requires time-consuming tasks in the design workflow. Our implementation is a first step towards a fully functional Digital Twin of a gerotor pump. Our implemented tool allows for the integration of data that are collected from an experimental setup with a virtual prototype model. The collected data are fed to the numerical model in order to improve the accuracy of the performance predictions. This allows the engineer to have a fast overview of the performance of the pump and allows him to discard unsuitable geometric configurations in an efficient manner. The presented implementation integrates a 2D design interface with an interactive parameterized model of the pump and a 3D interface. The 3D tool allows for the visualization of the 3D model that corresponds to the previously defined 2D geometry. Our initial tests show that the implemented model to perform fast pre-CFD simulations approaches the result of more detailed and time-consuming simulations within an acceptable margin of error. Our implemented tool also integrates, in a single application, the geometry data and simulation data, which are otherwise treated in different environments. Our virtual prototype is not suitable if a detailed prediction of the behavior inside each chamber of the pump is required.

Future work is needed in the improvement of both the software implemented and the fast simulation model in order to achieve a fully functional Digital Twin tool of the pump. Efforts regarding the 3D tool should be focused on data visualization and the creation of a data structure for simulation profiles of different geometric configurations. A needed improvement of the pre-CFD simulation may be achieved by modeling (a) cavitation and (b) micro-movements in the rotors, due to induced pressure at the chambers. A further integration of experimental data with the simulation model is needed to improve the accuracy of the predictions.

Author Contributions: J.P.-C., A.M., A.P.-B., E.S.-J. and O.R.-S. conceptualized the tool; J.P.-C., A.M. and B.S. implemented the tool; J.P., O.R.-S., E.S.-J. and A.P.-B. supervised the industrial application of the tool; A.M., B.S. and J.P.-C. implemented the geometrical aspects of this research; J.P.-C., E.S.-J. and A.P.-B. supervised the simulation model results. All the authors contributed to the writing of the article. All authors have read and agreed to the published version of the manuscript.

Funding: This work has received funding from the Eusko Jaurlaritza/Basque Government under the grants KK-2018/00071 (LANGILEOK) and ZL-2020/00190 (LATIDO).

Acknowledgments: The authors thank Tecnun-Universidad de Navarra for the support in the CFD-generated operational data.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

LP	Lumped Parameter.
CFD	Computational Fluid Dynamics.
P_i	Pressure at chamber *i*.
A_i	Area at chamber *i*.
V_i	Volume at chamber *i*.
Q_i	Net flowrate at chamber *i*.
$A_{i,out}$	Shared area between chamber *i* and output port.
$A_{i,in}$	Shared area between chamber *i* and input port.
C_d	Discharge coefficient.
$C_{d,max}$	Maximum discharge coefficient.
β_{eff}	Effective bulk's modulus of working fluid.
ρ_{eff}	Effective density of working fluid.
ν	Kinematic viscosity of working fluid.
μ	Dynamic viscosity of working fluid.
Re	Reynold's number.
Re_{crit}	Critical Reynold's number.
Dh	Hydraulic diameter.
p_{in}	Pressure at input port.
p_{out}	Pressure at output port.
ω	Angular speed of inner gear.

References

1. Gamez-Montero, P.J.; Codina, E.; Castilla, R. A review of gerotor technology in hydraulic machines. *Energies* **2019**, *12*, 2423. [CrossRef]
2. Castilla, R.; Gamez-Montero, P.; Ertürk, N.; Vernet, A.; Coussirat, M.; Codina, E. Numerical simulation of turbulent flow in the suction chamber of a gearpump using deforming mesh and mesh replacement. *Int. J. Mech. Sci.* **2010**, *52*, 1334–1342. [CrossRef]
3. Houzeaux, G.; Codina, R. A finite element method for the solution of rotary pumps. *Comput. Fluids* **2007**, *36*, 667–679. [CrossRef]
4. Hsieh, C.F. Fluid and dynamics analyses of a gerotor pump using various span angle designs. *J. Mech. Des.* **2012**, *134*. [CrossRef]
5. Bae, J.H.; Kwak, H.S.; San, S.; Kim, C. Design and CFD analysis of gerotor with multiple profiles (ellipse–involute–ellipse type and 3-ellipses type) using rotation and translation algorithm. *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2016**, *230*, 804–823. [CrossRef]
6. Rundo, M.; Altare, G. Lumped Parameter and Three-Dimensional CFD Simulation of a Variable Displacement Vane Pump for Engine Lubrication. *J. Fluids Eng.* **2018**, *140*, 61–101. [CrossRef]
7. Gamez-Montero, P.J.; Castilla, R.; del Campo, D.; Ertürk, N.; Raush, G.; Codina, E. Influence of the interteeth clearances on the flow ripple in a gerotor pump for engine lubrication. *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.* **2012**, *226*, 930–942. [CrossRef]
8. Pellegri, M.; Vacca, A.; Frosina, E.; Buono, D.; Senatore, A. Numerical analysis and experimental validation of Gerotor pumps: A comparison between a lumped parameter and a computational fluid dynamics-based approach. *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* **2017**, *231*, 4413–4430. [CrossRef]
9. Pellegri, M.; Vacca, A.; Devendran, R.S.; Dautry, E.; Ginsberg, B. A Lumped parameter approach for gerotor pumps: Model Formulation and experimental validation. In Proceedings of the 2016 10th International Fluid Power Conference, Dresden, Germany, 8–10 March 2016; Technische Universitat Dresden: Dresden, Germany, 2016; Volume 1, pp. 465–476.
10. Shah, Y.; Vacca, A.; Dabiri, S.; Frosina, E. A fast lumped parameter approach for the prediction of both aeration and cavitation in Gerotor pumps. *Meccanica* **2018**, *53*, 175–191. [CrossRef]
11. Rundo, M. Models for flow rate simulation in gear pumps: A review. *Energies* **2017**, *10*, 1261. [CrossRef]
12. Tao, F.; Zhang, H.; Liu, A.; Nee, A.Y. Digital twin in industry: State-of-the-art. *IEEE Trans. Ind. Inform.* **2018**, *15*, 2405–2415. [CrossRef]
13. Pires, F.; Cachada, A.; Barbosa, J.; Moreira, A.P.; Leitão, P. Digital Twin in Industry 4.0: Technologies, Applications and Challenges. In Proceedings of the 2019 IEEE 17th International Conference on Industrial Informatics (INDIN), Espoo, Finland, 22–25 July 2019; IEEE: Piscataway, NJ, USA, 2019; Volume 1, pp. 721–726.
14. Mejia, D.; Moreno, A.; Arbelaiz, A.; Posada, J.; Ruiz-Salguero, O.; Chopitea, R. Accelerated thermal simulation for three-dimensional interactive optimization of computer numeric control sheet metal laser cutting. *J. Manuf. Sci. Eng.* **2018**, *140*. [CrossRef]
15. Mejia-Parra, D.; Arbelaiz, A.; Ruiz-Salguero, O.; Lalinde-Pulido, J.; Moreno, A.; Posada, J. Fast Simulation of Laser Heating Processes on Thin Metal Plates with FFT Using CPU/GPU Hardware. *Appl. Sci.* **2020**, *10*, 3281. [CrossRef]
16. Gámez Montero, P.J. *Caracterizacion Fluidodinamica de una Bomba Oleohidráulica de Engranajes Internos Generados por Perfiles Trocoidales*; Universitat Politècnica de Catalunya: Barcelona, Spain, 2004.
17. Kwon, S.M.; Kang, H.S.; Shin, J.H. Rotor profile design in a hypogerotor pump. *J. Mech. Sci. Technol.* **2009**, *23*, 3459–3470. [CrossRef]
18. Kim, S.; Murrenhoff, H. Measurement of effective bulk modulus for hydraulic oil at low pressure. *J. Fluids Eng.* **2012**, *134*. [CrossRef]
19. Simões, B.; Creus, C.; Carretero, M.d.P.; Guinea Ochaíta, A. Streamlining XR Technology Into Industrial Training and Maintenance Processes. In *The 25th International Conference on 3D Web Technology*; Association for Computing Machinery: New York, NY, USA, 2020. [CrossRef]
20. Simões, B.; del Puy Carretero, M.; Santiago, J.M. Photorealism and Kinematics for Web-Based CAD Data. In *The 25th International Conference on 3D Web Technology*; Association for Computing Machinery: New York, NY, USA, 2020. [CrossRef]
21. Castilla López, R.; Gámez Montero, P.J.; Raush Alviach, G.A.; Codina Macià, E. Three dimensional simulation of gerotor with deforming mesh by using OpenFOAM. In Proceedings of the Fluid Power Networks: Proceedings: 19th-21th March 2018: 11th International Fluid Power Conference, Aachen, Germany, 19–21 March 2018; pp. 260–271.

applied sciences

Review

Is Digital Twin Technology Supporting Safety Management? A Bibliometric and Systematic Review

Giulio Paolo Agnusdei [1,2,*], Valerio Elia [1] and Maria Grazia Gnoni [1]

[1] Dipartimento di Ingegneria dell'Innovazione, Università del Salento, via per Monteroni, 73100 Lecce, Italy; valerio.elia@unisalento.it (V.E.); mariagrazia.gnoni@unisalento.it (M.G.G.)
[2] Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, S.P. Andersens vei, 5, NO-7491 Trondheim, Norway
* Correspondence: giulio.agnusdei@unisalento.it

Abstract: In the Industry 4.0 era, digital tools applied to production and manufacturing activities represent a challenge for companies. Digital Twin (DT) technology is based on the integration of different "traditional" tools, such as simulation modeling and sensors, and is aimed at increasing process performance. In DTs, simulation modeling allows for the building of a digital copy of real processes, which is dynamically updated through data derived from smart objects based on sensor technologies. The use of DT within manufacturing activities is constantly increasing, as DTs are being applied in different areas, from the design phase to the operational ones. This study aims to analyze existing fields of applications of DTs for supporting safety management processes in order to evaluate the current state of the art. A bibliometric review was carried out through VOSviewer to evaluate studies and applications of DTs in the engineering and computer science areas and to identify research clusters and future trends. Next, a bibliometric and systematic review was carried out to deepen the relation between the DT approach and safety issues. The findings highlight that in recent years, DT applications have been tested and developed to support operators during normal and emergency conditions and to enhance their abilities to control safety levels.

Keywords: safety; digital twin; smart operator; smart manufacturing; Industry 4.0

check for updates

Citation: Agnusdei, G.P.; Elia, V.; Gnoni, M.G. Is Digital Twin Technology Supporting Safety Management? A Bibliometric and Systematic Review. *Appl. Sci.* **2021**, *11*, 2767. https://doi.org/10.3390/app11062767

Academic Editor: A.Y.C. Nee

Received: 26 February 2021
Accepted: 18 March 2021
Published: 19 March 2021

1. Introduction

The fourth industrial revolution—also known as the Industry 4.0 paradigm—is changing current industrial production systems. The Industry 4.0 digital transformation is aimed at the optimization and automation of the previously introduced digitalization by adding the intelligent networking of machines, processes, and people [1]. In recent years, the concepts and theories of Industry 4.0 have become considerably important within the manufacturing sector [2]. One main reason for this increasing diffusion is the use and implementation of hi-tech technologies within the production processes at an affordable cost [3] in order to create smart grids across the entire value chain and networks formed by interconnecting intelligent machines [4]. Within industrial manufacturing contexts, Industry 4.0 entails the networking of data coming from machines, products, and people; in general, this involves the interconnection of smart devices among different plants and factories [5] through tools and embedded components, such as cyber–physical systems (CPSs), Internet of Things (IoT), cloud computing, robotics, systems based on artificial intelligence, and cognitive computation [4].

Digital Twins (DTs) are key enabling technologies (KETs) that were created to improve the efficiency and profitability of Industry 4.0 systems [6,7].

The first applications of DTs in industrial systems focused on developing a "digitalized copy" of a production process and/or a product. More recently, through the overlapping with the IoT, DT functionality and interoperability has been "augmented" by adding a real-time interaction with the real system [1].

The availability of virtual and digital information represents an effective solution for improving product design, manufacturing technology, and other critical service processes, such as safety management [8,9].

Safety management, as the process of realizing certain safety functions, aims to promote organizational safety and to protect people and property within an organization from unacceptable safety risks [10]. Recently, with the increasing global economic uncertainties, safety management in most organizations is under growing pressure to achieve the best performance, and this requires access to a variety of high-quality safety information [11].

In the era of Industry 4.0, it cannot be disputed that data and information represent an indispensable resource and a successful key factor for a new paradigm of safety management [12–15].

Within the Safety 4.0 framework [16], DT-assisted safety management systems can be implemented to help operators execute complex safety procedures, thus reducing risks and human errors. DT guidance can lead operators through safety tasks and provide them with real-time information regarding the contextual conditions [17]. This can reduce the costs and time required for service and maintenance, decrease oversights and mistakes, and increase safety [18,19].

The main purpose of this study is to present the prevailing state of research on Digital Twin technology, its manifold applications, and the intersection between DTs and safety, answering the following research questions:

RQ1. What are the current publication trends in the domain in terms of the types of studies, time, and affiliated countries?

RQ2. Which are the influential studies and themes of research in this domain and how have they evolved over the years?

RQ3. What are the recent research trends, gaps, and areas for future research in this domain?

The rest of this paper is organized as follows: Section 2 delineates the advent of DTs, their main features, and their background. Section 3 illustrates the adopted methodological framework, while Section 4 covers the findings on the publication trends and on the keyword and cluster content analyses. The study is concluded in Section 5, where future research avenues are suggested.

2. Background

The DT concept first appeared in the aerospace field around the 1970s (Figure 1); the first complete characterization was proposed by Grieves during a course on Product Lifecycle Management at the University of Michigan [20].

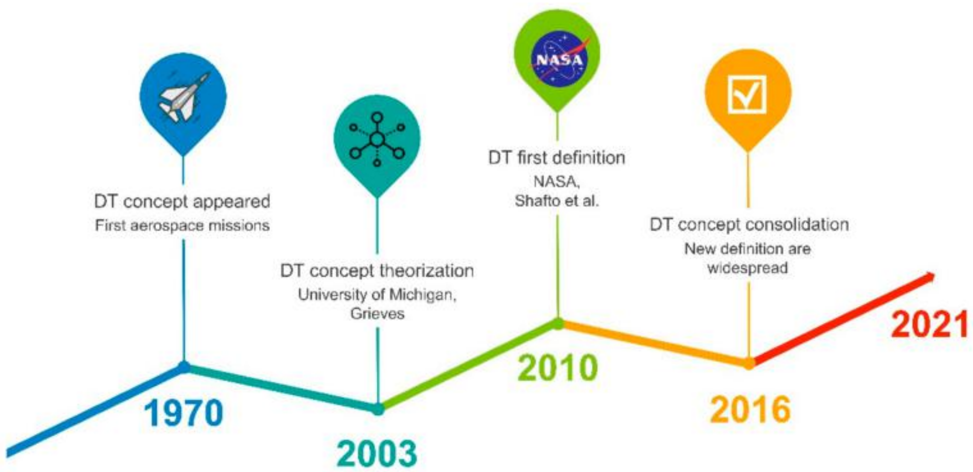


Figure 1. The timeline of the evolution of the definition of a Digital Twin (DT).

Basically, a DT refers to a system consisting of three main subsystems: (a) physical products in real space, (b) virtual products in virtual space, and (c) data and information that tie the virtual and real products together. Grieves [20] depicted DT flow as a cycle between the physical and virtual states (called twinning), with data flowing from the physical to the virtual and information and processes flowing from the virtual to the physical states (Figure 2).

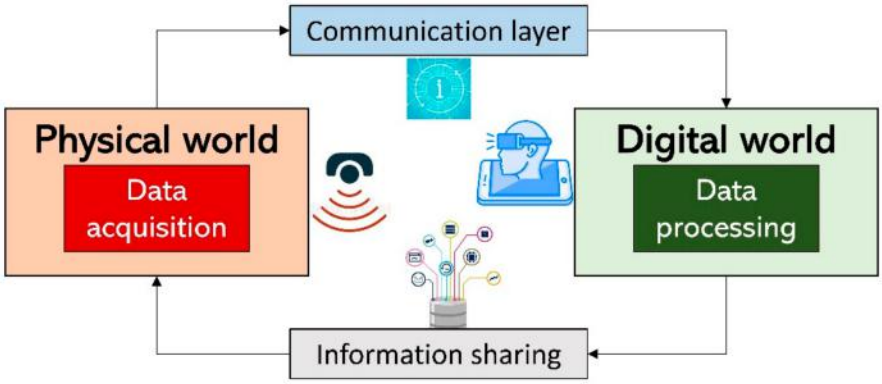


Figure 2. Digital Twin concept.

Then, NASA provided the first definition of a DT for the aeronautic sector [21]: A "digital twin is an integrated multi-physics, multi-scale, probabilistic simulation of a vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its flying twin".

Differently from the first definition, this one referred explicitly to the concept of simulation modeling, which is often adopted in several tools for supporting the design, validation, and testing of a system. Within this definition, the simulation is considered as a multi-physics model, as more than one physical field is simultaneously involved and different physical properties are integrated; simulation modeling is also usually multi-scale because different levels of time and/or space are used; finally, it is probabilistic because it could easily be developed with probability calculations.

These features characterize traditional simulation models, and they are enhanced in DT applications due to their stricter connection with the physical world.

Traditional simulation models are developed based on real data, but their updating is usually a static process [22]. Differently, in DTs, a data update process is required, and this represents the most relevant feature of DTs. Thus, due to the diffusion of the Industry 4.0 paradigm, a huge amount of data about physical systems are now available—even in real time—to be used for operation redesign and control. This availability is becoming the main key enabler of DT system development in manufacturing systems.

From 2010 onwards, the definition of DTs given by NASA was modified, and new definitions that were more focused on the industrial sector were proposed in the scientific literature. Lee et al. [23] proposed DTs as an advancement in predictive manufacturing systems. In their definition, the DT is defined as a simulation model that acquires real data and transfers them to a simulator in the cloud. In line with this study, Rosen et al. [24] stated that DTs, by combining real data with simulation models, allowed for the drawing up of a forecast based on realistic data, thus providing a sort of guidance system that supports operators and planners during normal operation, as well as during maintenance and service. Chen [25] described the DT as a computerized model of a physical device or system that represents all functional features and links with the working elements.

In 2018, some authors identified the Digital Twin as a digital representation of a physical production system that uses integrated simulations and service data, holding information from multiple sources across a product's life cycle. This information is continuously updated based on operational changes and is visualized in different ways to forecast current and future conditions of the physical counterpart in order to enhance decision making [26,27].

Recently, one of the latest definitions identified the Digital Twin as a virtual instance of a physical system (twin) that is continually updated with the latter's performance, maintenance, and health status data throughout the physical system's life cycle [28].

As a promising means of achieving cyber–physical interaction, integration, and fusion, Digital Twins (DTs) have captured growing attention from academic researchers as well as industrial practitioners [29,30].

They can stimulate the development of new approaches in design, production, and service, eventually leading to more innovations, such as better data management in order to improve the production process and performance and to ensure continuity and traceability of information [31], support for the analysis of production-line performance parameters, allowing for continuous monitoring of the line balancing and performance with the variations in the production demand [32], and support for monitoring and decision making regarding the ergonomic performance of manual production lines [33].

3. Methodology

Review studies can be of several types. In this study, a combination of bibliometric and systematic reviews is adopted. Bibliometric analyses are extensively performed to trace the knowledge anatomy of a research field and are used to analyze research topics [34]. Systematic literature reviews are used to synthesize the contents of the literature, limit bias [35], and identify possible research gaps.

For the purpose of answering RQ1 and RQ2 and identifying the publication trends and most influential research themes, two bibliometric analyses were performed, which provided comprehensive maps of the knowledge structures of the DT research field and the intersection between the DT and safety research fields. The results from the clustering of the intersection between the DT and safety research fields provided a foundation for a cluster content analysis aimed at answering RQ3 and identifying the recent research trends in the domain, as well as the gaps and areas for future research. Figure 3 proposes a schematization of the adopted methodology.

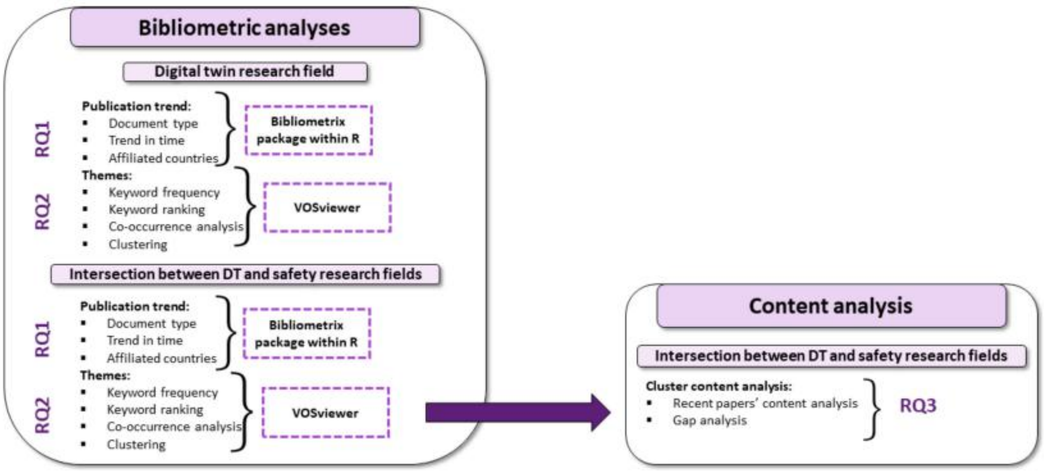


Figure 3. Methodological framework.

3.1. Search Protocol and Datasets

The Scopus database, the widest repository of peer-reviewed scientific literature, was used for the construction of the datasets. The first step was the identification of the keywords to be used for the selection of the document samples.

The terms selected to extract the document sample for the first bibliometric analysis (Figure 3) were the following: (i) digital; (ii) twin. Scopus was queried on 10 February

2021 with this combination of keywords: "digital AND twin". This choice allowed the borders of the analysis to be defined, ensuring a specific focus on the topics to be inspected. The extraction from the Scopus database was limited to: (i) studies written in the English language; (ii) the "Engineering" and "Computer science" areas; (iii) the period 2003–2021 because the first theorization of DTs was provided by Grieves in 2003. It provided an output of 3301 documents, whose citation information (authors, documents title, publication year), abstract, and index keywords were exported.

In order to limit the analysis to the intersection between the digital twin and safety management topics, a second extraction from the Scopus database was performed to obtain the document sample for the second bibliometric analysis (Figure 3). For the same time period, 2003–2021, Scopus was queried on 10 February, 2021 with the combination of keywords "digital AND twin AND safety", and with the same limitations as in the first bibliometric analysis. The obtained dataset consisted of 190 documents, whose citation information (authors, documents title, publication year), abstract, and index keywords were exported.

3.2. Bibliometric and Content Analysis Methods

The methodology adopted to investigate the scientific literature dynamics related to the two document samples extracted from Scopus was based on the use of two types of software for bibliometric analysis. The Bibliometrix package application in the R software was used to evaluate the growth, maps, and trends of the scientific field of research [36], while VOSviewer 1.6.14, which was developed to conduct text mining and to construct bibliometric maps [37], was used to identify the study keywords' co-occurrence.

In the study, the VOSviewer software was used to create a network map of the co-occurrence of terms considering index keywords. The keywords' co-occurrence analysis was carried out using the full counting method. The relatedness of items was determined based on the number of documents in which they occurred together.

According to the methods proposed by Donohue [38], the cut-off point for the term occurrence was determined with the following Formula (1):

$$T = \left(1 + \sqrt{1 + 8 \times I}\right)/2 \quad (1)$$

where T represents the optimal minimum number of occurrences of a keyword and I is the total number of keywords. Notwithstanding, to ensure wider software processing, keywords that co-occurred at least 10 times were selected for the analysis [39].

The index keywords were processed using VOSviewer, and the results were displayed in the network visualization and the overlay visualization. The network visualization shows keyword co-occurrence, where the dimensions of circles represent the weights of keywords, the lines represent the ways in which two words are linked, and thicker lines mean stronger connections among words. In the network visualization, VOSviewer uses colors to indicate the cluster to which a keyword has been assigned. The clustering technique [40] requires an algorithm for solving an optimization problem. For this purpose, VOSviewer implements the smart local moving algorithm introduced by Waltman and Van Eck [41].

The overlay visualization replicates the same map as that in the network visualization, but with different colors. The items and their links are colored in order to make it possible to view temporal trends and to identify which keywords were used most frequently during the observation period. The layout of the map was built by normalizing the strengths of the links among the elements through the association strength method according to Van Eck and Waltman [42].

In order to conduct a systematic review of the safety management issues related to the DT concept, documents were extracted from the clusters obtained through the second bibliometric phase. The most recent documents were carefully examined to identify the common features that characterized each cluster as well as the gaps in the literature.

4. Results

4.1. The Digital Twin Research Field

Conference papers represent essential documents for supporting the scientific development of new research and application fields. As shown in Figure 4, the analysis of the dataset according to the type of document shows that the majority are conference papers (57%), followed by journal articles (33%), conference reviews (4%), and book chapters (3%). These results represent important evidence that allows Digital Twin research to be identified as an emerging research field.

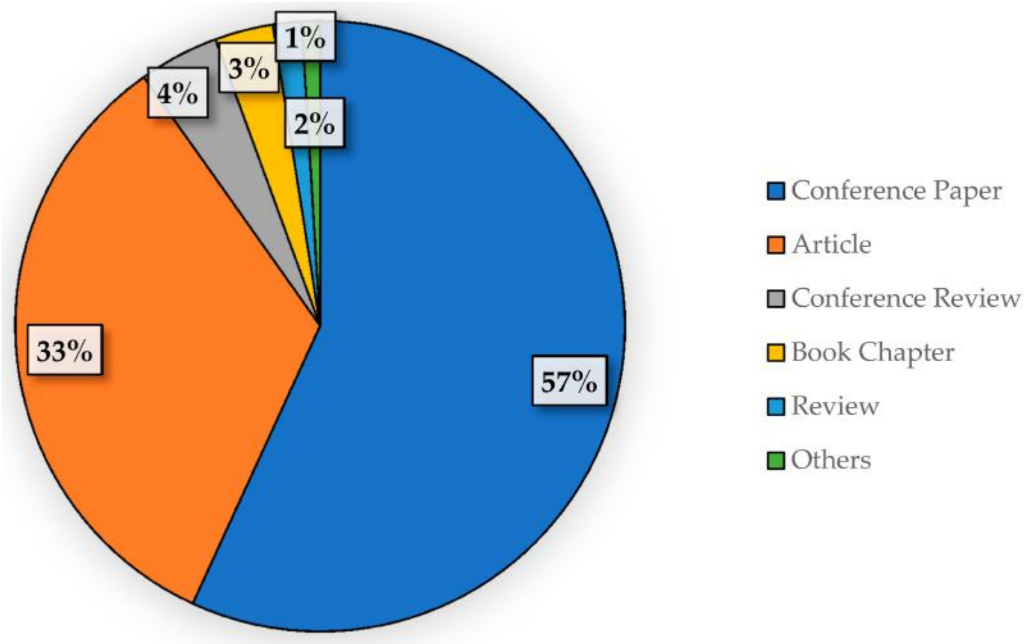


Figure 4. Publications in the DT research field by document type.

Figure 5 illustrates the progression of publications available in the Scopus data on Digital Twins in the period 2003–2021. There was an upsurge in publications from just 25 documents published in 2003 to 1279 documents in 2020. Research on Digital Twins saw a sudden spurt from 2017, which could represent a point of discontinuity, with an exponential shape in the last three years (2018–2020). The main reason for this surge may mainly be attributed to the governmental ICT (Information and Communication Technologies) investments in many countries carried out by the adoption of Industry 4.0 policy plans [43]. Even if there are more conference papers than journal articles in absolute terms, in the last three years (2018–2020), the share of journal articles over the total published documents regarding DTs has increased noticeably, going from 24% to 35%, registering an opposite trend compared to conference papers, which fell from 67% to 52%.

Table 1 lists the top countries affiliated with authors of DT research, with the leading three being Germany (492 documents), the United States (477 documents), and China (434 documents). They are the top three global manufacturing export countries in the world [44], and Germany, in addition to having a leading position in Europe (Figure 6), is the country where the Industry 4.0 concept was first developed [44,45]. A noticeable self-perpetuating effect of giving and taking references was observed in these countries, which register more study contributions to the pool of the DT field. Countries that registered a high production were also among the countries whose studies received more citations. It can be said that the most well-established countries are dominating or leading in the field, and this can also be seen in the citation patterns.

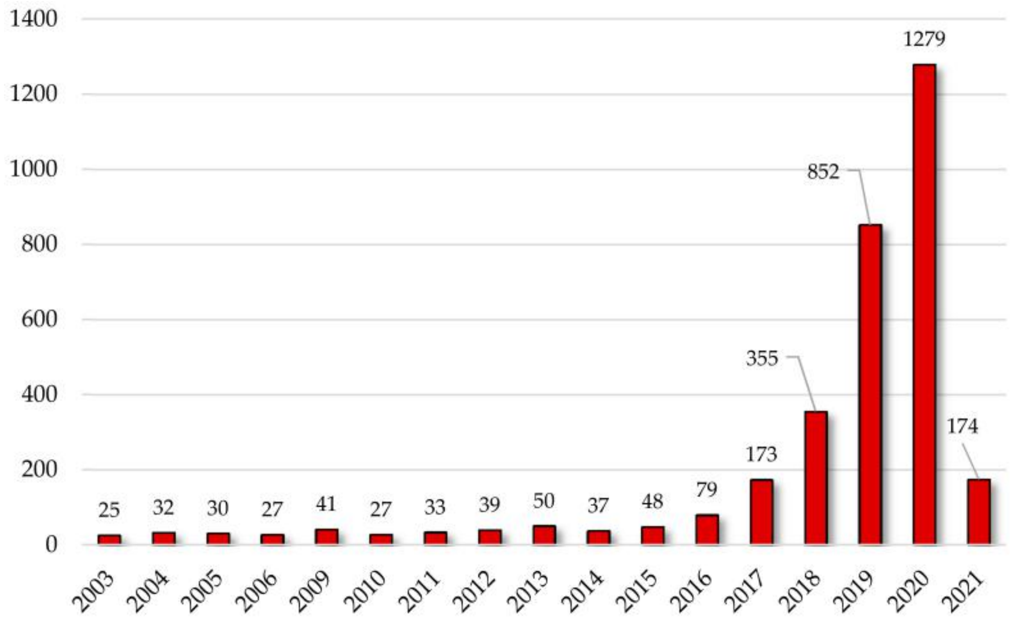


Figure 5. Annual publication trend of 3301 documents retrieved from Scopus for the DT research field in the period 2003–2021.

Table 1. Top ten affiliated countries publishing on DTs.

Country	Number of Documents
Germany	492
United States	477
China	434
United Kingdom	241
Russian Federation	205
Italy	164
France	144
India	116
Spain	95
South Korea	89

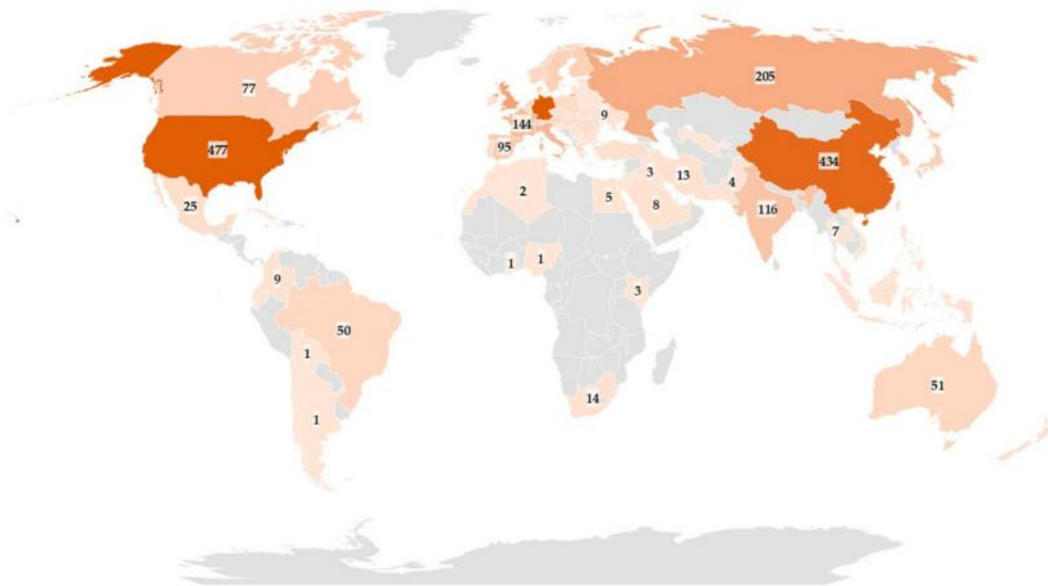


Figure 6. Map of affiliated countries publishing on DTs.

Keyword analysis was performed to explore the most prevalent themes in the DT research field. A total of 15,901 keywords were identified in 3301 documents. It is important to evaluate the keywords of a document to understand how authors frame their work, what the most interesting aspect of the research is or how it is evolving, and what trends are being created. From the extracted document sample, a count of the index keywords was performed in order to calculate their frequency and rank them. Table 2 shows the ranking of the top ten most relevant Keywords-Plus (ID). The ID is standardized; it is defined by Scopus to help in the research of documents associated with a topic. "Digital twin" is the most frequently used keyword, with 1404 occurrences, which indicates that this word alone is used as a termed concept in the literature. The other most frequently used keywords are "life cycle" (333 occurrences), "manufacture" (325 occurrences), "embedded systems" (248 occurrences), "Internet of Things" (206 occurrences), and "Industry 4.0" (200 occurrences).

Table 2. Top ten most frequent index keywords in the DT research field.

Keyword	Frequency
Digital twin	1404
Life cycle	333
Manufacture	325
Embedded systems	248
Internet of things	206
Industry 4.0	200
Decision making	179
Cyber–physical system	169
Virtual reality	153
Digital storage	151

Based on the hypothesis that a research specialty can be identified by the relations among document keywords, keyword co-occurrence analysis is useful for identifying the thematic areas or clusters that constitute the theoretical blocks or foundational topics of the field under analysis [46].

Starting from the entire document sample, including 15,901 keywords, a co-occurring keyword analysis was performed in order to construct diagrams to display the network and the overlay visualization (respectively, Figures 7 and 8). Considering keywords that co-occurred at least 10 times, 502 keywords were selected for the final analysis.

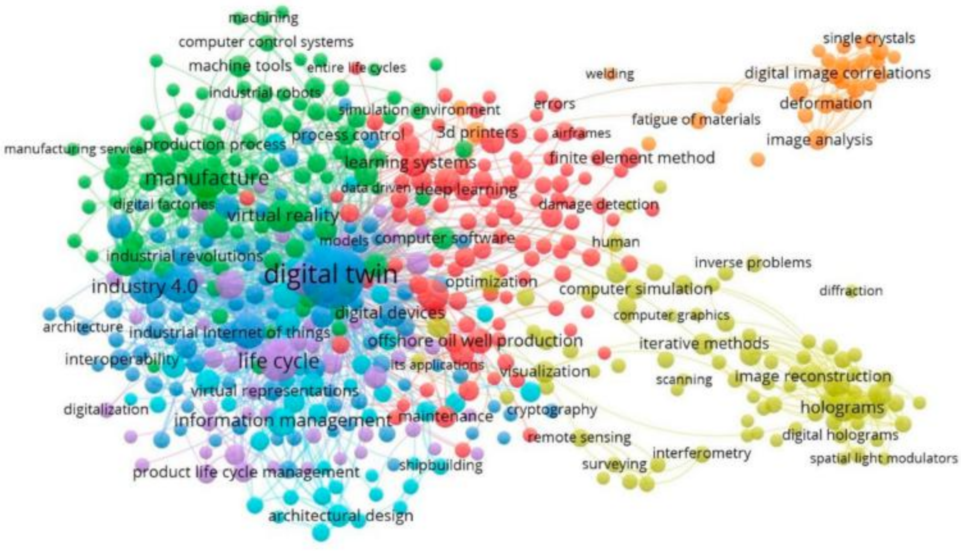


Figure 7. Network visualization of the DT research field.

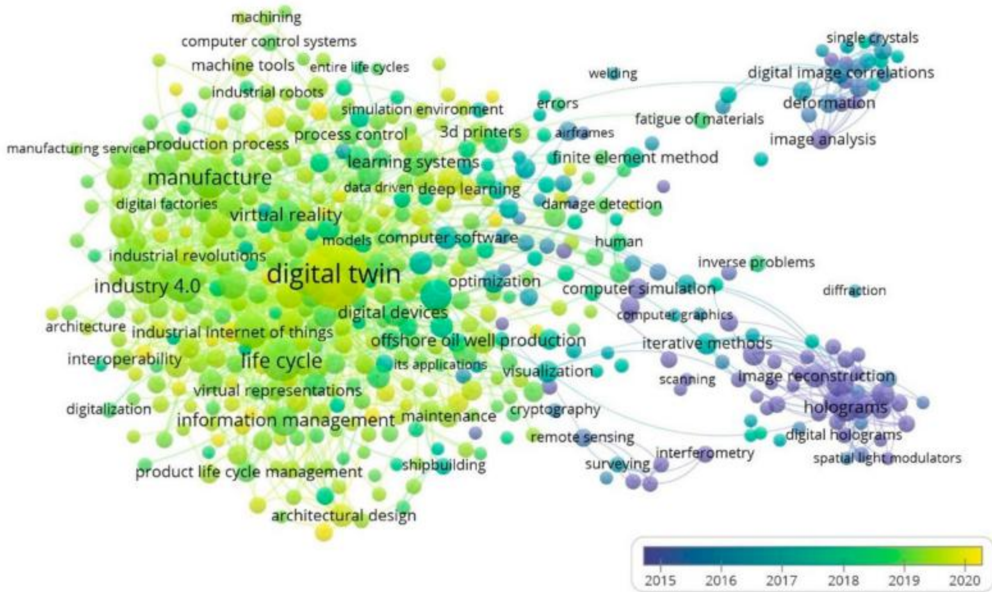


Figure 8. Overlay visualization of the DT research field.

The keywords were grouped into clusters, represented by different colors. In particular, the keywords were clustered into seven groups, and each keyword was assigned to only one cluster. As shown by the network visualization (Figure 7), the DT research field is primarily composed of seven clusters of connected topics: Industry 4.0 and IoT (cluster—blue), learning systems (cluster—red), smart manufacturing (cluster—green), information management (cluster—cyan), life-cycle management (cluster—purple), holograms (cluster—yellow), and digital image analysis (cluster—orange).

A very close linkage between five clusters (red, blue, green, purple, and cyan) is clearly observable, while two clusters appear disconnected from the first ones (yellow and orange).

The clusters clearly indicate that scientific research in the field of DTs has focused, above all, on Industry 4.0 manufacturing, life-cycle management, and data processing and analysis. Since it is evident that the orange and yellow clusters are almost totally unlinked from each other, the holograms and digital image analysis research fields can be considered to have developed almost autonomously.

In the overlay visualization, which shows the temporal distribution of the keywords in each cluster (Figure 8), keywords are colored according to a score. This score is given based on the average year of occurrence of a keyword. Colors range from blue (oldest time period) to green and yellow (most recent time periods). It emerged that the topics related to holograms (cluster—yellow) and digital image analysis (cluster—orange) were developed before the others. This could mean that in an early stage, the DT concept and research field derived previously from holograms, digital image analysis, and scanning, and then evolved to a wider and more common application that is usually adopted in manufacturing activities.

The most recent documents in the scientific literature refer, in fact, to the fields of smart manufacturing, IoT, life cycle, and information management (Figure 8), which also represent the main research fields according to the co-occurrence network map (Figure 7). The latter are trendy topics. Within them, there are a series of strongly linked research sub-fields, e.g., quality control of processes and products, design of processes and products, predictive analysis through machine learning algorithms, and, finally, safety management.

4.2. Bibliometric Results on the Intersection between the DT and Safety Research Fields

Safety aspects are fundamental within the manufacturing sector and represent a research field with great potential and wide application areas [47]. As indicated in Section 3, the linkages between the Digital Twin and safety issues were investigated through a second bibliometric analysis.

As for the document sample related to Digital Twins, the document sample related to studies that integrate the DT and safety research fields also includes many conference proceedings. The latter generally report preliminary studies, which serve as a basis for the development of more complex research activities that later conclude with international journal publication [48]. Figure 9 shows that conference papers represent 65% of the dataset according to the type of document, followed by journal articles (33%), conference reviews (7%), reviews (4%), and book chapters (1%). In the light of these results, studies that integrate DTs and safety issues can be considered as the most recent emerging research field within the wider research field of DTs [49].

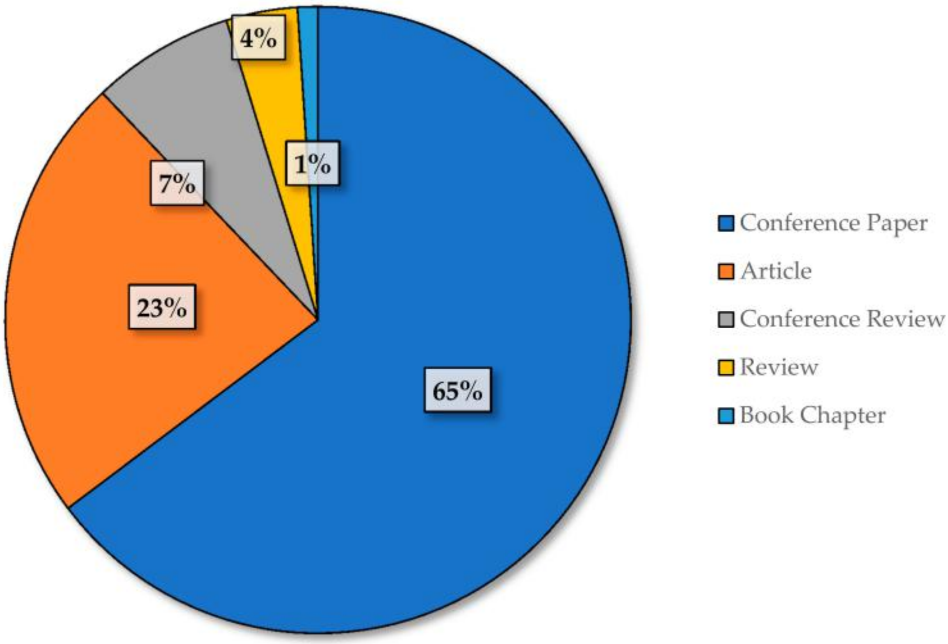


Figure 9. Publications in the intersection of the DT and safety research fields by document type.

Figure 10 highlights the trend of publications available in the Scopus data on the intersection between the Digital Twin and safety research fields in the period 2003–2021. Most of the articles were published in the three-year period of 2018–2020. In fact, there was, an exponential increase in publications from 0 documents published in 2003 to 92 documents in 2020, which represents a peak, confirming the researchers' pioneering interest in developing the scientific topics of Digital Twins related to process manufacturing activities and, consequently, to safety management.

Table 3 lists the top countries affiliated with the authors of research on the intersection of DTs and safety, with the leading three being the United States (40 documents), Germany (25 documents), and China (21 documents), confirming the rank results registered for the DT research field, with the only permutation being between Germany and the United States. The numbers of documents published by authors from the UK (16 documents) and Italy (13 documents) are also noteworthy. Italy earned the fifth position, also highlighting the role that the national "Industry 4.0" plan played in driving organizational changes in enterprises that particularly addressed safety issues [50]. Figure 11 shows Map of affiliated countries publishing in the intersection of the DT and safety research fields. Even if they are not very numerous, documents from Italy register a high impact on the research field. As an example, the review conducted by Cimino et al. [51] (2019) of the Politecnico di Milano, which was published in *Computers in Industry* and dealt with a topic related to the intersection between DTs and safety, is among the most cited documents (43 citations). This highly cited document is followed by the study by Oyekan et al. [52] (34 citations), which shows how virtual reality Digital Twins could assist in the safe implementation of human–robot collaborative strategies in factories of the future.

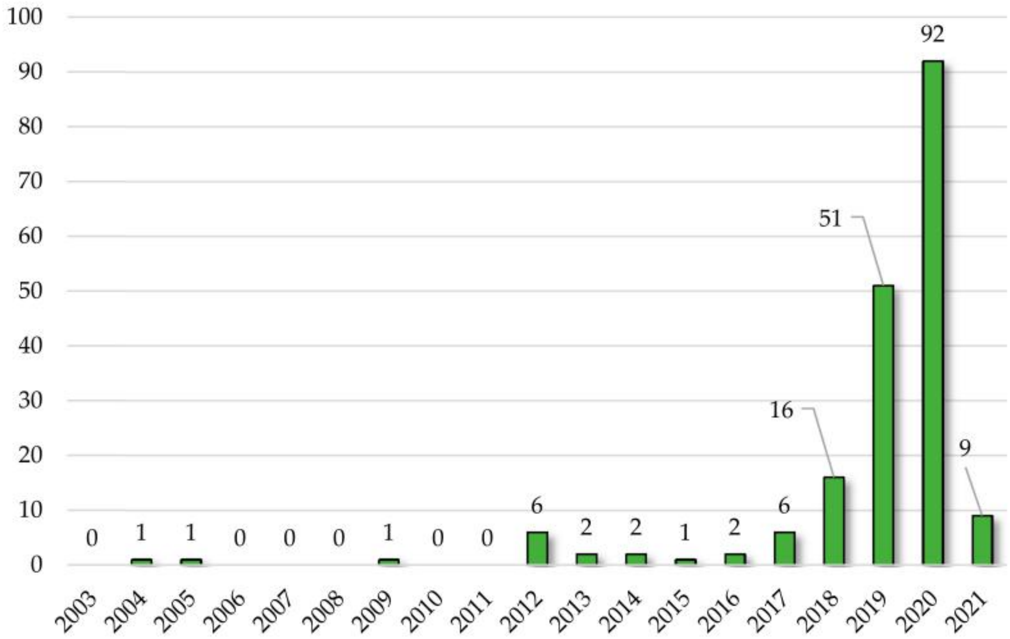


Figure 10. Annual publication trend of 3301 documents retrieved from Scopus for the intersection between the DT and safety research fields in the period 2003–2021.

Table 3. Top ten affiliated countries publishing in the intersection of the DT and safety research fields.

Country	Number of Documents
United States	40
Germany	25
China	21
United Kingdom	16
Italy	13
Russian Federation	10
Austria	9
Sweden	8
India	7
Norway	6

Keyword analysis was performed to explore the most prevalent themes in the intersection between the DT and safety research fields. A total of 1673 keywords were identified in 190 documents. From the extracted document sample, a count of the index keywords was performed in order to calculate their frequency and to rank them. Table 4 shows the ranking of the top ten most relevant Keywords-Plus (ID). As for the first bibliometric analysis, "Digital twin" is the most frequently used keyword, with 93 occurrences, followed by "life cycle" (32 occurrences), "safety engineering" (21 occurrences), "accident prevention" (19 occurrences), and "virtual reality" (19 occurrences).

A significant finding outlined by the keyword analysis is that there are themes that occur in both the DT research field and the intersection of the DT and safety research fields. This circumstance highlights the unanimity on the conceptualization of the Digital Twin and its applications, as well as the mostly indirect and diversified references to safety aspects in much of the scientific literature on DTs.

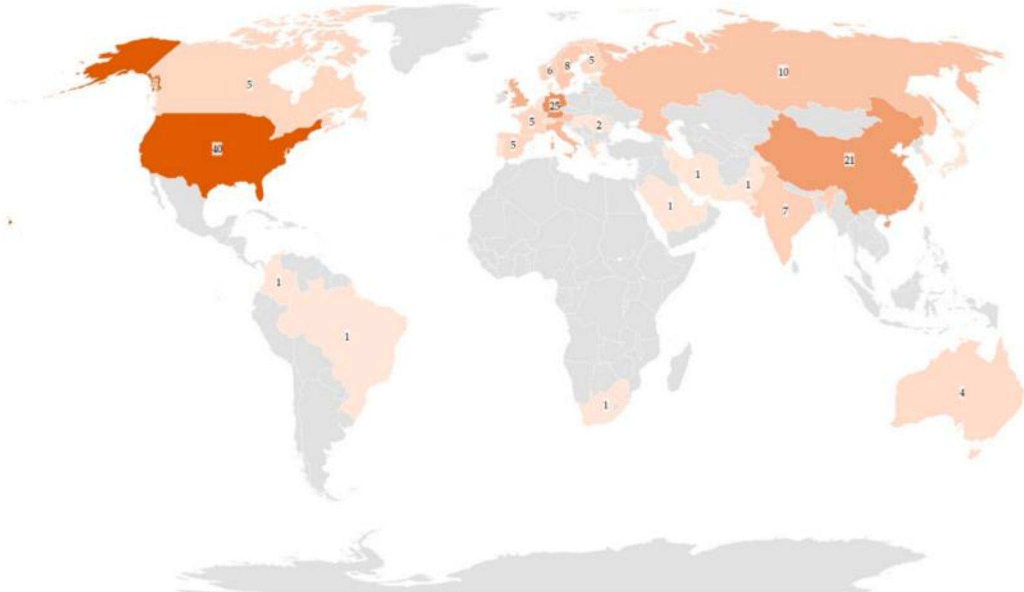


Figure 11. Map of affiliated countries publishing in the intersection of the DT and safety research fields.

Table 4. Top ten most frequent index keywords in intersection of the DT and safety research fields.

Keyword	Frequency
Digital twin	93
Life cycle	32
Safety engineering	21
Accident prevention	19
Virtual reality	19
Internet of Things	17
Embedded systems	17
Manufacture	16
Offshore oil well production	12
Offshore technology	10

Starting from the document sample, which included 1673 keywords, a co-occurring keyword analysis was performed in order to obtain Figures 12 and 13, which display the network and the overlay visualization for the intersection between the DT and safety research fields, respectively. In this case, 51 keywords were selected for the analysis, since they co-occurred at least 10 times.

As shown in Figure 12, the research field derived from the intersection between DTs and safety consists of seven clusters: decision making and offshore applications (cluster—red), IoT and life-cycle approaches (cluster—green), Industry 4.0: from manufacture to virtual reality (cluster—blue), machine learning support for DTs and safety (cluster—yellow), safety engineering (cluster—purple), hazards and risk assessment (cluster—orange), and DTs in battery management systems (cluster—cyan).

In the overlay visualization, which shows the temporal distribution of the keywords in each cluster (Figure 13), the field of studies regarding the intersection between DTs and safety has evolved from a previous concentration on topics related to aircraft and fleet operations to wider issues related to accident prevention and risk and information management, as well as to more strategic and specific themes referring to life-cycle management and offshore technology applications.

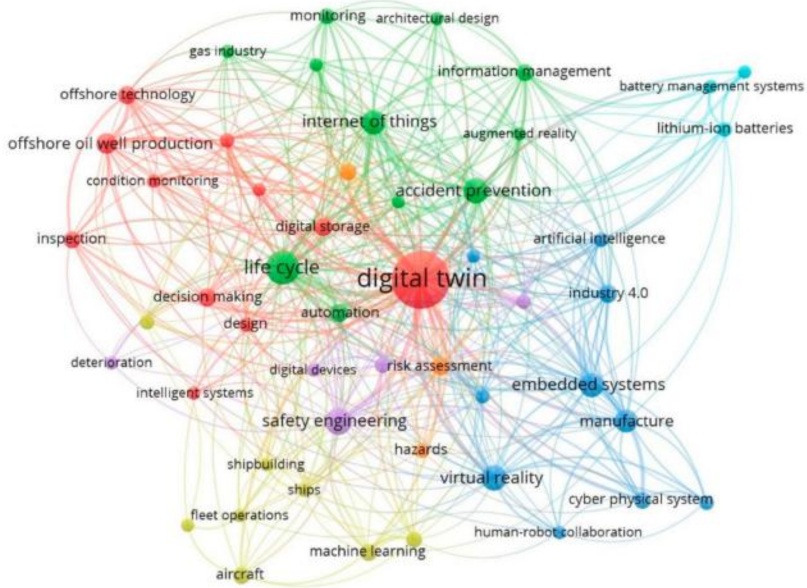


Figure 12. Network visualization of the intersection of the DT and safety research fields.

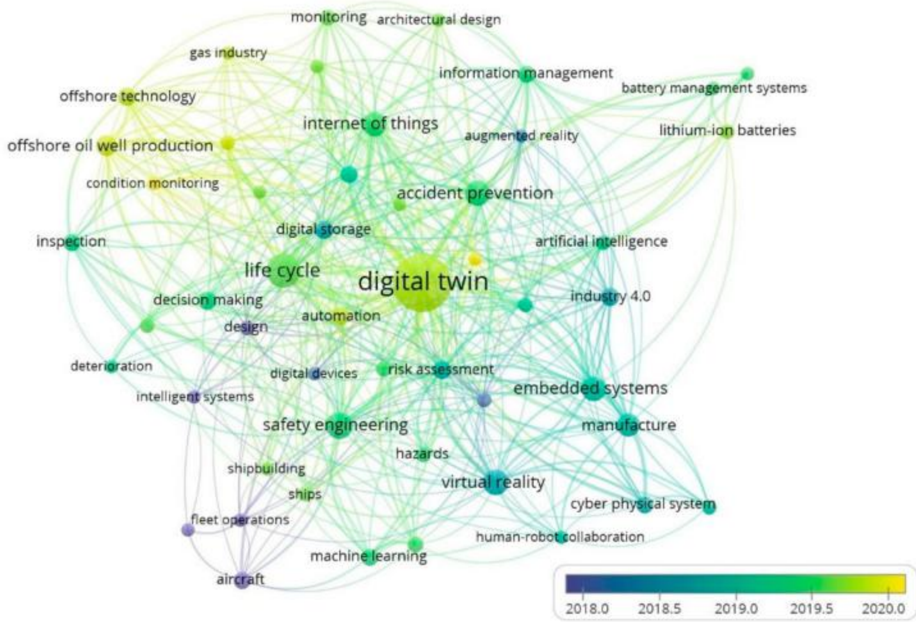


Figure 13. Overlay visualization of the intersection of the DT and safety research fields.

4.3. Cluster Content Results for the Intersection between the DT and Safety Research Fields

As shown in Figure 12, VOSviewer uses colors to indicate the cluster to which a keyword has been assigned within the network visualization based on the methods explained in Section 3.2.

The red cluster of keywords, called "Decision making and offshore applications", includes studies regarding DT solutions aimed at supporting decision-making processes. These studies identify the Digital Twin concept as a promising tool for decision makers and stakeholders alike, which is bound to benefit those who use it [53]. In particular, DTs solve big data problems in the field of offshore resources, letting workers spend less time looking for data and more time identifying trends and innovative ways to exploit the data, e.g., smarter drilling, greater field automation, or improved safety [54]. Based on in situ measurement information, DTs can support operational and maintenance decisions that will preserve the integrity, safety, and availability of assets [55].

The green cluster, called "IoT and life-cycle approaches", regards studies that consider that a DT should encompass and plan the entire life cycle of a physical asset, thus producing profound differences depending on the application domain [56]. Through the use of the IoT, the gap between the physical and virtual worlds is filled by bridging a physical component's sensors and actuators with its digital counterpart [57]. In the safety domain, a DT provides an opportunity to train employees in virtual environments, thus helping to achieve accident prevention and to reduce the probability of accidents that may occur during on-the-job training [58].

The blue cluster, called "Industry 4.0: from manufacture to virtual reality", groups the latest research related to Digital Twins and virtual reality environments for safety purposes within the Industry 4.0 paradigm. Over recent years, the concept of human–machine interaction has received wide attention, since it represents the basis for achieving automation in manufacturing. Conventional simulations do not allow us to experience future production systems as end-users in an immersive environment. For this reason, virtual reality has had technological development [59]. Cyber–physical systems, however, require operators' awareness of the situation in order to be able to adequately address potential issues in a timely manner. Detecting early symptoms may speed up the incident response process and mitigate the consequences of business interruption or safety hazards. Running parallel to their physical counterparts, DTs allow for the deep inspection of their behavior without the risk of disrupting operational technology processes [60].

The yellow cluster, called "Machine learning support for DT and safety", includes studies that use machine learning algorithms to rapidly ascertain optimal aircraft dynamics to maximize the fire-retardant release effectiveness [61], to create a digital representation of humans that focuses on their vital quantities, as well as on the surrounding environment, for health monitoring purposes [62], or to enable industrial robots to bypass obstacles or people in a workspace [63].

The purple cluster, called "Safety engineering", regards studies about the interactions among DT data acquisition, DT data processing, and safety issues [49], as well as research regarding advanced structural simulations combined with physics-based deterioration models in order to calculate structural performance [64].

The orange cluster, called "Hazards and risk assessment", groups studies about risk assessment and tools of risk prediction. By comparing the real data with those obtained by the simulation software, DTs can predict risks and/or anomalies and communicate with a server in order to generate a warning [65]. This is particularly relevant when DTs serve to audit and evaluate compliance with legal requirements in everyday production and logistics processes. Non-compliance can endanger employees and the environment, and can cause financial and reputational damage [66].

The cyan cluster, called "DT in battery management systems", refers to studies that aim to enhance the safety, reliability, and performance of battery systems. All data relevant to batteries can be measured and transmitted to create a Digital Twin of a battery system, allowing for the diagnostic evaluation of a battery's charge and aging level [67].

As highlighted through the cluster content analysis, the scientific literature lacks a generic and widely recognized DT architecture. This heterogeneity unavoidably requires a deeper understanding of how to deal with DT systems by evaluating and comparing them and exploring how they differ in handling different environments. The lack of standardization and the variety of DT definitions, which cause discrepancies among DT implementation projects, represent a challenge that requires further research efforts aimed at fastening the progress in supporting safety management.

When trying to address open research questions in the field of Digital Twins, another challenge comes from the multidisciplinary approaches needed to design and develop adequate safety management measures. With technological improvements and the spread of blockchain, the scientific gap related to the integration of data for small IoT systems, as well as large heterogeneous systems, including human interaction, should be filled.

5. Conclusions

The Digital Twin technology is an emerging topic that has captured the attention of researchers in recent years. It is also becoming popular amongst managers and practitioners, thus demonstrating that it is one of the most fertile and contributing fields in the areas of engineering and computer science research.

This review further contributes to DT research in terms of unfolding the evolving literature in terms various themes and trends, hence putting forth the status of scholarly work since its inception in 2003. However, some gaps in the research on the intersection between the DT and safety research fields have been identified, and recent research efforts were addressed.

With the intention of thoroughly reviewing the existing literature, this study provides valuable insights into DTs and safety. Such concepts have an ever-increasing importance in day-to-day management decisions. The scholarly work reveals the still scarce implementation of DT technologies across industry, and specifically for safety management purposes. This study represents a wake-up call for decision makers and other stakeholders who should undertake paths towards improvements enabled by DT technologies and, ultimately, safety in its complex definition. There is a large scope of contributions to theoretical development, methodologies, and new applications. DT technology is an issue with vast implications for safety management, and its development can guide the way to competitive and stable industries.

Author Contributions: Conceptualization, G.P.A. and M.G.G.; Data curation, G.P.A.; Methodology, G.P.A., M.G.G., and V.E.; Supervision, M.G.G.; Validation, G.P.A.; Visualization, G.P.A.; Writing—original draft, M.G.G., G.P.A., and V.E.; Writing—review and editing, M.G.G. and V.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research is part of the activities carried out within SO4SIMS project (Smart Operators 4.0 based on Simulation for Industry and Manufacturing Systems) funded by the Italian Ministry of Education, Universities and Research MIUR (Project PRIN-2017FW8BB4).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jacoby, M.; Usländer, T. Digital Twin and Internet of Things—Current Standards Landscape. *Appl. Sci.* **2020**, *10*, 6519. [CrossRef]
2. Culot, G.; Nassimbeni, G.; Orzes, G.; Sartor, M. Behind the definition of Industry 4.0. Analysis and open questions. *Int. J. Prod. Econ.* **2020**, *226*, 107617. [CrossRef]
3. Turel, M.; Akis, E. Industry 4.0 and Competitiveness. *Res. J. Bus. Manag.* **2019**, *6*, 204–212. [CrossRef]
4. Florescu, A.; Barabaș, S.A. Modeling and Simulation of a Flexible Manufacturing System—A Basic Component of Industry 4.0. *Appl. Sci.* **2020**, *10*, 8300. [CrossRef]
5. Lezzi, M.; Lazoi, M.; Corallo, A. Cybersecurity for Industry 4.0 in the current literature: A reference framework. *Comput. Ind.* **2018**, *103*, 97–110. [CrossRef]
6. Negri, E.; Fumagalli, L.; Macchi, M. A review of the roles of digital twin in CPS-based production systems. *Procedia Manuf.* **2017**, *11*, 939–948. [CrossRef]
7. De Miranda, S.S.-F.; Aguayo-González, F.; Salguero-Gómez, J.; Ávila-Gutiérrez, M.J. Life cycle engineering 4.0: A proposal to conceive manufacturing systems for industry 4.0 centred on the human factor (DfHFinI4. 0). *Appl. Sci.* **2020**, *10*, 4442. [CrossRef]
8. Wang, X.; Ong, S.K.; Nee, A.Y.C. A comprehensive survey of augmented reality assembly research. *Adv. Manuf.* **2016**, *4*, 1–22. [CrossRef]
9. Khalek, I.A.; Chalhoub, J.M.; Ayer, S.K. Augmented reality for identifying maintainability concerns during design. *Adv. Civ. Eng.* **2019**, *3*, 8547928. [CrossRef]
10. Li, Y.; Guldenmund, F.W. Safety management systems: A broad overview of the literature. *Saf. Sci.* **2018**, *103*, 94–123. [CrossRef]
11. Huang, L.; Wu, C.; Wang, B.; Ouyang, Q. Big-data-driven safety decision-making: A conceptual framework and its influencing factors. *Saf. Sci.* **2018**, *109*, 46–56. [CrossRef]
12. Gobbo, J.A., Jr.; Busso, C.M.; Gobbo, S.C.O.; Carreão, H. Making the links among environmental protection, process safety, and industry 4.0. *Process Saf. Environ. Prot.* **2018**, *117*, 372–382. [CrossRef]

13. Huang, L.; Wu, C.; Wang, B. Challenges, opportunities and paradigm of applying big data to production safety management: From a theoretical perspective. *J. Clean. Prod.* **2019**, *231*, 592–599. [CrossRef]
14. Wang, B.; Wu, C.; Huang, L.; Kang, L. Using data-driven safety decision-making to realize smart safety management in the era of big data: A theoretical perspective on basic questions and their answers. *J. Clean. Prod.* **2019**, *210*, 1595–1604. [CrossRef]
15. Lee, J.; Cameron, I.; Hassall, M. Improving process safety: What roles for Digitalization and Industry 4.0? *Process Saf. Environ. Prot.* **2019**, *132*, 325–339. [CrossRef]
16. Wang, B. Safety intelligence as an essential perspective for safety management in the era of Safety 4.0: From a theoretical to a practical framework. *Process Saf. Environ. Prot.* **2021**, *148*, 189–199. [CrossRef]
17. Gattullo, M.; Scurati, G.W.; Evangelista, A.; Ferrise, F.; Fiorentino, M.; Uva, A.E. Informing the use of visual assets in industrial augmented reality. In Proceedings of the International Conference on Design Tools and Methods in Industrial Engineering, Modena, Italy, 9–10 September 2020.
18. Martinettia, A.; Rajabalinejada, M.; van Dongena, L. Shaping the future maintenance operations: Reflections on the adoptions of augmented reality through problems and opportunities. *Procedia CIRP* **2017**, *59*, 4–17. [CrossRef]
19. Siew, C.Y.; Ong, S.K.; Nee, A.Y.C. Improving maintenance efficiency and safety through a human-centric approach. *Adv. Manuf.* **2021**, *9*, 104–114. [CrossRef]
20. Grieves, M.W. Product lifecycle management: The new paradigm for enterprises. *Int. J. Prod. Dev.* **2005**, *2*, 71–84. [CrossRef]
21. Shafto, M.; Conroy, M.; Doyle, R.; Glaessgen, E.; Kemp, C.; LeMoigne, J.; Wang, L. *Modeling, Simulation, Information Technology & Processing Roadmap*; National Aeronautics and Space Administration: Washington, DC, USA, 2012.
22. Boschert, S.; Rosen, R. Digital twin—The simulation aspect. In *Mechatronic Futures*; Hehenberger, P., Bradley, D., Eds.; Springer: Cham, Switzerland, 2016; pp. 59–74. [CrossRef]
23. Lee, J.; Lapira, E.; Bagheri, B.; Kao, H.A. Recent advances and trends in predictive manufacturing systems in big data environment. *Manuf. Lett.* **2013**, *1*, 38–41. [CrossRef]
24. Rosen, R.; Von Wichert, G.; Lo, G.; Bettenhausen, K.D. About the importance of autonomy and digital twins for the future of manufacturing. *IFAC PapersOnLine* **2015**, *48*, 567–572. [CrossRef]
25. Chen, Y. Integrated and intelligent manufacturing: Perspectives and enablers. *Engineering* **2017**, *3*, 588–595. [CrossRef]
26. Zheng, Y.; Yang, S.; Cheng, H. An application framework of digital twin and its case study. *J. Ambient Intell. Humaniz. Comput.* **2019**, *10*, 1141–1153. [CrossRef]
27. Vrabič, R.; Erkoyuncu, J.A.; Butala, P.; Roy, R. Digital twins: Understanding the added value of integrated models for through-life engineering services. *Procedia Manuf.* **2018**, *16*, 139–146. [CrossRef]
28. Madni, A.M.; Madni, C.C.; Lucero, S.D. Leveraging digital twin technology in model-based systems engineering. *Systems* **2019**, *7*, 7. [CrossRef]
29. Tao, F.; Qi, Q.; Wang, L.; Nee, A.Y.C. Digital twins and cyber–physical systems toward smart manufacturing and industry 4.0: Correlation and comparison. *Engineering* **2019**, *5*, 653–661. [CrossRef]
30. Tao, F.; Liu, A.; Hu, T.; Nee, A.Y.C. *Digital Twin Driven Smart Design*; Academic Press: Cambridge, MA, USA, 2020.
31. Pang, T.Y.; Restrepo, J.D.P.; Cheng, C.T.; Yasin, A.; Lim, H.; Miletic, M. Developing a digital twin and digital thread framework for an 'Industry 4.0' Shipyard. *Appl. Sci.* **2021**, *11*, 1097. [CrossRef]
32. Fera, M.; Greco, A.; Caterino, M.; Gerbino, S.; Caputo, F.; Macchiaroli, R.; D'Amato, E. Towards digital twin implementation for assessing production line performance and balancing. *Sensors* **2020**, *20*, 97. [CrossRef]
33. Greco, A.; Caterino, M.; Fera, M.; Gerbino, S. Digital Twin for Monitoring Ergonomics during Manufacturing Production. *Appl. Sci.* **2020**, *10*, 7758. [CrossRef]
34. Merigó, J.M.; Blanco-Mesa, F.; Gil-Lafuente, A.M.; Yager, R.R. Thirty years of the International Journal of Intelligent Systems: A bibliometric review. *Int. J. Intell. Syst.* **2017**, *32*, 526–554. [CrossRef]
35. Tranfield, D.; Denyer, D.; Smart, P. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *Br. J. Manag.* **2003**, *14*, 207–222. [CrossRef]
36. Aria, M.; Cuccurullo, C. Bibliometrix: An R-tool for comprehensive science mapping analysis. *J. Informetr.* **2017**, *11*, 959–975. [CrossRef]
37. Van Eck, N.J.; Waltman, L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **2010**, *84*, 523–538. [CrossRef] [PubMed]
38. Donohue, J.C. *Understanding Scientific Literature: A Bibliographic Approach*; The MIT Press: Cambridge, UK, 1973.
39. Valente, D.; Miglietta, P.P.; Porrini, D.; Pasimeni, M.R.; Zurlini, G.; Petrosillo, I. A first analysis on the need to integrate ecological aspects into financial insurance. *Ecol. Model.* **2019**, *392*, 117–127. [CrossRef]
40. Waltman, L.; Van Eck, N.J.; Noyons, E.C. A unified approach to mapping and clustering of bibliometric networks. *J. Informetr.* **2010**, *4*, 629–635. [CrossRef]
41. Waltman, L.; Van Eck, N.J. A smart local moving algorithm for large-scale modularity-based community detection. *Eur. Phys. J. B* **2013**, *86*, 1–14. [CrossRef]
42. Van Eck, N.J.; Waltman, L. How to normalize cooccurrence data? An analysis of some well-known similarity measures. *J. Am. Soc. Inf. Sci. Technol.* **2009**, *60*, 1635–1651. [CrossRef]
43. Sung, T.K. Industry 4.0: A Korea perspective. *Technol. Forecast. Soc. Chang.* **2018**, *132*, 40–45. [CrossRef]

44. Kuo, C.C.; Shyu, J.Z.; Ding, K. Industrial revitalization via industry 4.0—A comparative policy analysis among China, Germany and the USA. *Global Transit.* **2019**, *1*, 3–14. [CrossRef]
45. Rojko, A. Industry 4.0 concept: Background and overview. *Int. J. Interact. Mob. Technol.* **2017**, *11*, 77–90. [CrossRef]
46. Manesh, M.F.; Pellegrini, M.M.; Marzi, G.; Dabic, M. Knowledge management in the fourth industrial revolution: Mapping the literature and scoping future avenues. *IEEE Trans. Eng. Manag.* **2020**, *68*, 289–300. [CrossRef]
47. Pichard, R.; Philippot, A.; Saddem, R.; Riera, B. Safety of manufacturing systems controllers by logical constraints with safety filter. *IEEE Trans. Control Syst. Technol.* **2018**, *27*, 1659–1667. [CrossRef]
48. Paez, A. Gray literature: An important resource in systematic reviews. *J. Evid. Based Med.* **2017**, *10*, 233–240. [CrossRef] [PubMed]
49. Agnusdei, G.P.; Elia, V.; Gnoni, M.G. A classification proposal of digital twin applications in the safety domain. *Comput. Ind. Eng.* **2021**, 107137. [CrossRef]
50. Federmeccanica. Final Report INDUSTRY 4EU—Industry 4.0 for the Future of Manufacturing in Europe. Available online: http://adapt.it/Industry4EU/INDUSTRY%204EU_final_report.pdf (accessed on 24 February 2021).
51. Cimino, C.; Negri, E.; Fumagalli, L. Review of digital twin applications in manufacturing. *Comput. Ind.* **2019**, *113*, 103130. [CrossRef]
52. Oyekan, J.O.; Hutabarat, W.; Tiwari, A.; Grech, R.; Aung, M.H.; Mariani, M.P.; López-Dávalos, L.; Ricaud, T.; Singh, S.; Dupuis, C. The effectiveness of virtual environments in developing collaborative strategies between industrial robots and humans. *Robot. Comput. Integr. Manuf.* **2019**, *55*, 41–54. [CrossRef]
53. Nåfors, D.; Berglund, J.; Gong, L.; Johansson, B.; Sandberg, T.; Birberg, J. Application of a Hybrid Digital Twin Concept for Factory Layout Planning. *Smart Sustain. Manuf. Syst.* **2020**, *4*, 231–244. [CrossRef]
54. Brewer, T.; Knight, D.; Noiray, G.; Naik, H. Digital twin technology in the field reclaims offshore resources. In Proceedings of the Offshore Technology Conference, Houston, TX, USA, 6–9 May 2019.
55. Anderson, S.; Barvik, S.; Rabitoy, C. Innovative digital inspection methods. In Proceedings of the Offshore Technology Conference, Houston, TX, USA, 6–9 May 2019.
56. Boje, C.; Guerriero, A.; Kubicki, S.; Rezgui, Y. Towards a semantic Construction Digital Twin: Directions for future research. *Autom. Constr.* **2020**, *114*, 103179. [CrossRef]
57. Yuan, X.; Anumba, C.J. Cyber-Physical Systems for Temporary Structures Monitoring. In *Cyber-Physical Systems in the Built Environment*; Anumba, C.J., Roofigari-Esfahan, N., Eds.; Springer: Cham, Switzerland, 2020; pp. 107–138. [CrossRef]
58. Wanasinghe, T.R.; Wroblewski, L.; Petersen, B.; Gosine, R.G.; James, L.A.; De Silva, O.; Mann, G.K.I.; Warrian, P.J. Digital twin for the oil and gas industry: Overview, research trends, opportunities, and challenges. *IEEE Access* **2020**, *8*, 104175–104197. [CrossRef]
59. Malik, A.A.; Masood, T.; Bilberg, A. Virtual reality in manufacturing: Immersive and collaborative artificial-reality in design of human-robot workspace. *Int. J. Comput. Integr. Manuf.* **2020**, *33*, 22–37. [CrossRef]
60. Eckhart, M.; Ekelhart, A.; Weippl, E. Enhancing cyber situational awareness for cyber-physical systems through digital twins. In Proceedings of the 2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Zaragoza, Spain, 10–13 September 2019; pp. 1222–1225.
61. Zohdi, T.I. A digital twin framework for machine learning optimization of aerial firefighting and pilot safety. *Comput. Methods Appl. Mech. Eng.* **2021**, *373*, 113446. [CrossRef]
62. Scheuermann, C.; Binderberger, T.; von Frankenberg, N.; Werner, A. Digital twin: A machine learning approach to predict individual stress levels in extreme environments. In Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers, Cancùn, Mexico, 12–16 September 2020; pp. 657–664.
63. Dröder, K.; Bobka, P.; Germann, T.; Gabriel, F.; Dietrich, F. A machine learning-enhanced digital twin approach for human-robot-collaboration. *Procedia CIRP* **2018**, *76*, 187–192. [CrossRef]
64. Zandi, K.; Ransom, E.H.; Topac, T.; Chen, R.; Beniwal, S.; Blomfors, M.; Shu, J.; Chang, F.-K. A Framework for Digital Twin of Civil Infrastructure—Challenges & Opportunities. *Struct. Health Monit.* **2019**. [CrossRef]
65. Bottani, E.; Vignali, G.; Tancredi, G.P.C. A digital twin model of a pasteurization system for food beverages: Tools and architecture. In Proceedings of the 2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Cardiff, UK, 15–17 June 2020; pp. 1–8.
66. Perez, G.C.; Korth, B. Digital Twin for Legal Requirements in Production and Logistics based on the Example of the Storage of Hazardous Substances. In Proceedings of the 2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Singapore, 14–17 December 2020; pp. 1093–1097.
67. Li, W.; Rentemeister, M.; Badeda, J.; Jöst, D.; Schulte, D.; Sauer, D.U. Digital twin for battery systems: Cloud battery management system with online state-of-charge and state-of-health estimation. *J. Energy Storage* **2020**, *30*, 101557. [CrossRef]

applied sciences

Article

Sustainability Requirements of Digital Twin-Based Systems: A Meta Systematic Literature Review

Rui Carvalho [1,*] and Alberto Rodrigues da Silva [2]

[1] INOV.PT, Instituto Superior Técnico, Universidade de Lisboa, 1000-029 Lisboa, Portugal
[2] INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, 1000-029 Lisboa, Portugal; alberto.silva@tecnico.ulisboa.pt
* Correspondence: ruimiguelc@tecnico.ulisboa.pt

Abstract: Sustainable development was defined by the UN in 1987 as development that meets the needs of the present without compromising the ability of future generations to meet their own needs, and this is a core concept in this paper. This work acknowledges the three dimensions of sustainability, i.e., economic, social, and environmental, but its focus is on this last one. A digital twin (DT) is frequently described as a physical entity with a virtual counterpart, and the data, connections between the two, implying the existence of connectors and blocks for efficient and effective data communication. This paper provides a meta systematic literature review (SLR) (i.e., an SLR of SLRs) regarding the sustainability requirements of DT-based systems. Numerous papers on the subject of DT were also selected because they cited the analyzed SLRs and were considered relevant to the purposes of this research. From the selection and analysis of 29 papers, several limitations and challenges were identified: the perceived benefits of DTs are not clearly understood; DTs across the product life cycle or the DT life cycle are not sufficiently studied; it is not clear how DTs can contribute to reducing costs or supporting decision-making; technical implementation of DTs must be improved and better integrated in the context of the IoT; the level of fidelity of DTs is not entirely evaluated in terms of their parameters, accuracy, and level of abstraction; and the ownership of data stored within DTs should be better understood. Furthermore, from our research, it was not possible to find a paper discussing DTs only in regard to environmental sustainability.

Keywords: digital twins (DTs); Internet of Things (IoT); sustainability requirements; sustainable development; product design

check for updates

Citation: Carvalho, R.; da Silva, A.R. Sustainability Requirements of Digital Twin-Based Systems: A Meta Systematic Literature Review. *Appl. Sci.* **2021**, *11*, 5519. https://doi.org/10.3390/app11125519

Academic Editor: Andrew Y. C. Nee

Received: 31 March 2021
Accepted: 4 June 2021
Published: 15 June 2021

1. Introduction

A digital twin (DT) is often described as a digital or virtual entity with a physical counterpart, and with data connections between the two [1], implying the existence of connectors and blocks to allow efficient and effective data communication. A DT is a digital representation of some more complex physical system, and in spite of the fact there are distinct definitions of DT, this was the original, and the one we adopt [1]. Grieves and Vickers of NASA are considered the pioneers of this concept, presenting it in a lecture on product life-cycle management in 2003, as is acknowledged by Liu et al. [2]. They point out three components [2]: (i) a physical product, (ii) a virtual representation of that product, and (iii) the bi-directional data connections from the physical to the virtual representation, or vice versa. Among the main purposes of developing DTs are product design, modeling, simulation, and optimization of specific assets [3,4].

Today the usage of DTs is not yet generalized, but since 2015 there has been a clear increase of scientific studies toward a better understanding of their potentialities. Machine tools and consumer goods are common examples of DT usage. However, it is not in all cases that a DT has to be a high-fidelity digital model of a physical system or asset. For instance, a DT can also be used for representing a whole city (urban digital twin), geographic areas, buildings, or even human bodies and human organs. However, the focus of this paper is on DTs as they represent physical assets. Other reported usages of building DTs are for cybersecurity incident prediction [3], monitoring ergonomics in IoT contexts [5], online education [6], or optimization farming systems [7].

Product design is also a fundamental and related aspect of DTs and environmental sustainability because it is a discipline that deals with many complex decisions and cross-cutting concerns, such as safety, security, usability, or sustainability (including the choice of materials or the use of energy). Furthermore, product design may influence the planning of a production line, another frequent application of DTs [1,2,8]. This means that errors and failures can be predicted and managed with the help of DT approaches, using data analytics, artificial intelligence (AI), and machine-learning techniques.

Sustainable development is defined by the UN (1987), in the Brundtland Report ("Our Common Future") [9], as development that meets the needs of the present without compromising the ability of future generations to meet their own needs. This framework is very useful, as the concept of a circular economy (CE) [10], defined as an economic model to minimize the consumption of finite resources, is becoming more important and is closely related to the concept of the supply chain [11]. Industry 4.0 was introduced in 2011, and it has become synonymous with smart manufacturing/factory, corresponding to established concepts such as computer-integrated manufacturing, a flexible manufacturing system [10], CE, and all together, the management of a huge amount of data generated by DTs [12], allowing the integration of new tools such as those related to the Internet of Things (IoT).

Closely linked to product design optimization and CE is the concept of sustainable product design, defined by Massey [13] as the art of designing buildings, cities, and other artifacts so that they meet the objectives of sustainable development. Product design is not just art, it is also in the decision of what materials to use, choosing in such a way that a product is useful for society. A DT, as defined in this research, has a close connection with sustainable product design, but our focus will be on its IT characteristics and benefits.

This paper provides a meta systematic literature review (i.e., an SLR on SLRs) on the topic of the sustainability requirements of DT-based systems. In this context, sustainability requirements are defined as requirements that make sustainable development possible. Merten et al. [14] tried using the knowledge of generic requirements to assist automatically during requirement specification. Paech et al. [15] tried developing a systematic process for deriving the sustainability requirements for a specific system, i.e., a checklist of general and IT-specific details for each sustainability dimension (environmental, technical, social, economic, and individual), and the influences between them. Additionally, a new model and new concepts, such as needs and the effects between needs, whether negative, neutral, or positive, were created; this new mindset is used to develop a system respecting a due balance between different dimensions to achieve sustainability. In the context of IT, this means a controlled natural language [16] that helps the specification of requirements and tests systematically and rigorously [17,18], such as the ITLingo RSL language [19], does have a paramount role in defining software sustainability requirements [20] as it already supports risks, vulnerabilities and goals/solutions [21], for example. These tools belong to the spectrum of model-driven engineering, using textual specifications and conceptual models to improve the efficacy and efficiency of the analysis and design of these IT systems [22] and their usability [23].

To sum up, this paper is a meta-SRL plus an attempt to add new contributions to the environmental sustainability debate. Sustainable development implies a responsible consumption of resources today, and because DTs can allow optimization of the operations, they can be a tool for that purpose. When designers test a new product, they might use a DT to virtually test a new implementation without consuming raw materials and simulating the usage of environmentally friendly materials, also reducing working hours, and only produce it afterward if the simulation makes sense.

This paper is organized as follows: Section 2 introduces the SLR methodology followed; Section 3 presents the results of the meta-SLR; Section 4 presents a critical analysis to identify future work paths. Finally, Section 5 presents the main conclusions.

2. The Research Methodology

To develop this study, we considered the SLR methodology as proposed by Kitchenham et al. for software engineering [24], then we looked to the work by Escallón and Aldea because they presented a methodology that is valuable in the context of this research [25].

The method of SLR, as proposed by Kitchenham et al. [24], has three main stages: a phase for planning, to execute that plan, and to analyze the results. The execution phase has five tasks: (i) extract studies from databases, (ii) eliminate duplicates from the sample, (iii) apply inclusion and exclusion criteria, (iv) gather backward and forward citations, and (v) identify the final dataset of selected papers. If, at task iv, there are new papers found, then the researcher moves backward to task (ii) and repeats the process, from task (ii) to task (iv), as many times as are needed. Additionally, we included certain techniques described by Wolfswinkel et al., namely, the backward and forward citation steps of their selection phase [26]. If any work related to a research task is not present in the currently selected set of references, but it is considered relevant, it should be added to the selected ones. Since Google Scholar is a very popular tool and supports the backward and forward analysis of citations, this was the main tool adopted. The overall process is shown in Figure 1, inspired by the methodology used by Escallón and Aldea [25].

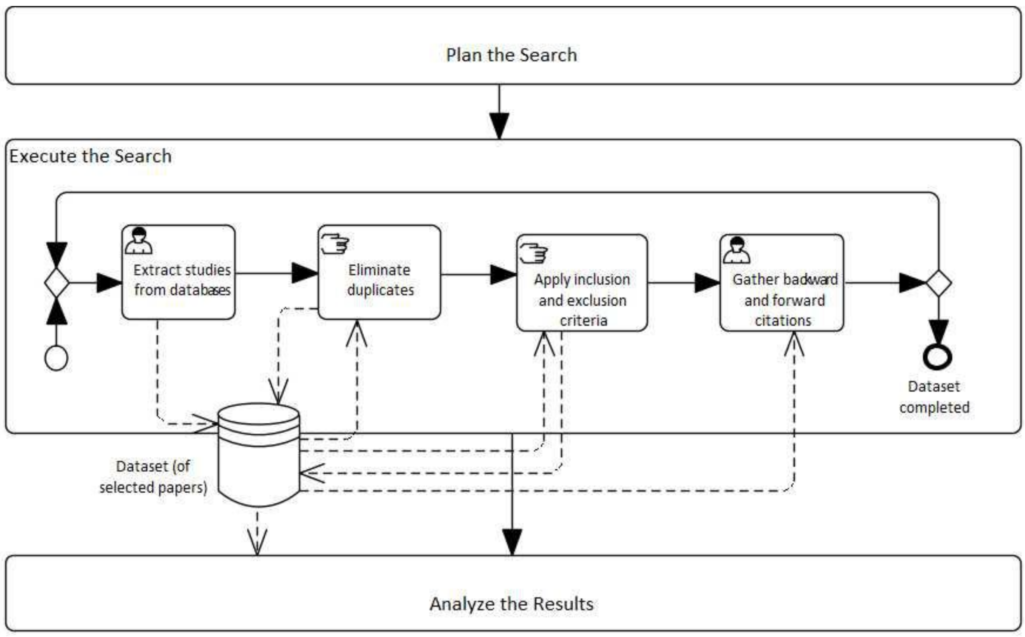


Figure 1. The overall process of SLR (in BPMN notation).

To allow for a more precise fine-tuning of our research, we also considered the work by Ahmad et al. [27]. They explicitly identified and categorized the different types of controlled or common vocabularies (CV) [27] available, and usage for the requirements specification of software development, which was an important output to us because we wished to focus on identifying sustainability requirements. A CV is an organized collection of terms that have well-known meanings, without the ambiguities or misunderstandings that synonyms could cause. Their purpose is to organize information in a structured manner with consistency, indicating semantic relationships, allowing the simple classification, querying, and retrieving of data [13,28–30]. Examples of the most frequent CVs are ontologies, taxonomies, thesauri, and folksonomies. Natural language processing and knowledge management techniques [31,32] often use CV support tools.

2.1. Planning Phase

The planning phase is the part where the design process is followed to perform the SLR according to the selected methodology. Firstly, the relationship between sustainability requirements and DTs-based systems is addressed, and then the relationship with product design is established. Section 2.1.1 defines the main questions that guided this research; Sections 2.1.2 and 2.1.3 refer to the scientific repositories and queries that were used in the search process, as well as the inclusion and exclusion criteria.

2.1.1. Research Question

We defined one research question (RQ1) with three sub-questions (SQ). These questions were used to address the main objectives of the research:

1. RQ1: What is the state of the art in the area of sustainability requirements of DT-based systems related to product design?
2. SQ1: What is the relationship between DTs and product design?
3. SQ2: What are the environmental sustainability requirements of DTs?
4. SQ3: What are the open issues and challenges in future research paths for DTs and sustainability?

The search process considers these questions and the technical definitions available in the existing literature and subsequent reading of the found works is also needed.

2.1.2. Search Process

The following were the databases or scientific repositories to search for relevant papers to answer the questions of this research. These databases were selected because they are very well known among the scientific community, and were also used in other SLR papers that we considered as models to follow in the IT domain:

1. ACM Digital Library (dl.acm.org (accessed on 4 June 2021))
2. Taylor & Francis (www.tandfonline.com (accessed on 4 June 2021))
3. Web of Science (www.webofknowledge.com (accessed on 4 June 2021))
4. ScienceDirect/Scopus (www.scopus.com (accessed on 4 June 2021))
5. IEEE Explore (ieeexplore.ieee.org/search/advanced (accessed on 4 June 2021))
6. Elsevier (www.elsevier.com (accessed on 4 June 2021))
7. Google Scholar (scholar.google.com (accessed on 4 June 2021))

These were the inclusion criteria: (i) articles published in the past 10 years; (ii) studies published in journals and conference proceedings or indexed books; (iii) studies written in English; (iv) articles referring to SLR that were selected in the first database searches. We also included high-quality studies, even when they were short or written in Portuguese. Conversely, these were the exclusion criteria: (i) very short papers (i.e., with fewer than 5 pages); and (ii) duplicated works, those unified by the database under various names.

2.1.3. Queries

In the first week of January 2021, we started from the query "systematic literature review" AND "digital twin" (all fields), and then tried other queries, considering not only the main aim of this paper but also the number of papers found, and the possibility of making the search deeper. The first step was to search for papers regarding a meta-SLR about DTs, and the second step was to search for sustainability requirements. The time period of the search was also considered, firstly the publication dates between 01/01/2011 and 31/12/2020, and secondly, the publication dates between 01/01/2019 and 31/12/2020. As a consequence, several additional queries were used:

- ACM: [All: "systematic literature review"] AND [All: "requirements"] AND [All: "digital twin*"] AND [Publication Date: (01/01/2019 TO 12/31/2020)];
- Elsevier: "digital twins" AND "systematic literature review" AND "product design" AND "requirement" AND "sustainability";

- IEEE: ((“Document Title”:”digital twin*”) AND (“All Metadata”:”systematic literature review”) AND (“All Metadata”:”design”)) and ((“Document Title”:”ontolo*”) AND (“All Metadata”:”systematic literature review”) AND (“All Metadata”:”design”));
- SCOPUS: TITLE-ABS-KEY (“systematic literature review” AND “digital twin*”) and TITLE-ABS-KEY (“systematic literature review” AND “ontolo*” AND “digital twin”);
- Web of Science: [All: “digital twin”] AND [All: “systematic literature review”] AND [All: “product design”] (Tailor); TITLE: (“systematic literature review” AND “digital twin*”) and TITLE: (“systematic literature review” AND “ontolo*”).

In different databases, we elicited distinct scientific outputs, and consequently, we had to use other queries. The search for the keyword “environmental sustainability” was fruitless.

2.2. Execution Phase

The execution phase is the step where both the results and the process to execute the SLR are explained. We followed the phases and criteria previously defined, and will now describe our experience during this process, followed by the useful information we were able to extract. The queries were presented in the already mentioned databases, and the results are presented in Table 1. After obtaining the results from these databases, we completed the following steps: first, eliminate all duplicates; second, based on the paper title, whether exclusion criteria apply; third, based on the title, select those articles where both inclusion criteria apply, and exclusion criteria do not apply; fourth, repeat the third step but read through the full text; fifth, for each remaining article, review the reference section and repeat steps 2 to 4; and sixth, for relevant references, for each remaining article, Google Scholar is used to review the forward citations and repeat steps 2 to 4.

Table 1. Search in Databases.

Database	First Search	With Google Scholar	Selected Papers	
			SLR	Non-SLR
ACM	6	15	3	7
ELSEVIER	19	10	1	2
IEEE Explore	13	34	1	4
SCOPUS	12	43	4	1
TAILOR	4	41	2	2
WEB SCIENCE	7	16	2	0
TOTAL	61	159	13	16

A first search (see column “First Search” in Table 1) allowed us to identify works related to SLR, and a second search identified citations using Google Scholar (see column “with Google Scholar”). It was also possible to identify relevant papers to analyze the context of DT usage even though they were not SLRs (see column “Non-SLR”). Finally, reading abstracts of the papers, it was possible to select 29 papers (column “Selected Papers”, with 13 SLR and 17 non-SLR papers). SLR papers are listed in Table 2, and non-SLR papers are listed in Table 3. The types of papers considered are C—Conference Paper; J—Journal Paper, T—Thesis, and B—Book.

Selected SLR papers are very recent: 2020 (10 papers); 2019 (2 papers) and 2018 (1 paper). The same holds true for selected non-SLR papers: 2020 (12 papers); 2019 (2 papers); 2016 (1 paper) and 2013 (1 paper). This situation is not surprising because the technologies and issues surveyed in this meta-SLR are very recent.

Table 2. Set of selected SLR papers.

ID	Reference	Title	Year	Type	Topics
S1	[1]	Characterising the Digital Twin: A systematic literature review	2020	J	DTs Definition
S2	[2]	Review of digital twin about concepts, technologies, and industrial applications	2020	J	DTs Definition
S3	[3]	Digital Twin for Cybersecurity Incident Prediction: A Multivocal Literature Review	2020	C	DTs Cybersecurity
S4	[4]	Reference Framework for Digital Twins within Cyber-Physical Systems	2019	C	DTs Cyber-Physical Systems
S5	[10]	Assessing relations between Circular Economy and Industry 4.0: a systematic literature review	2020	J	DTs, CE, I4.0
S6	[33]	Impact of Industry 4.0 on Sustainability—Bibliometric Literature Review	2020	J	DTs I4.0
S7	[34]	Industry 4.0 Model for circular economy and cleaner production	2020	J	DTs, CE I4.0
S8	[35]	Systematization of Digital Twins: Ontology and Conceptual Framework	2020	C	DTs Ontology
S9	[36]	A Review of the Literature on Smart Factory Implementation	2019	C	DTs, I4.0
S10	[37]	Digital Twins: Current problems in Smart City and Recommendations for future technology	2020	J	DTs Smart City
S11	[38]	Digital twin of stone sawing processes	2020	J	DTs, Example
S12	[39]	A Systematic Literature Review on the Application of Ontologies in Automatic Programming	2018	J	DTs Ontologies
S13	[40]	The Digital Twin Concept in Industry—A Review and Systematization	2020	J	DTs

Table 3. Set of selected non-SLR papers.

ID	Reference	Title	Year	Type	Topics
NS1	[8]	Enhancing Operational Performance and Productivity Benefits by Implementing Smart Manufacturing Technologies in Breweries	2019	J	I4.0
NS2	[41]	A review of industry 4.0 potential to accelerate the transition to a circular economy	2020	T	I4.0, CE
NS3	[42]	IoT-Based Digital Twin for Energy Cyber-Physical Systems: Design and Implementation	2020	J	DTs, Energy
NS4	[43]	A Framework for Quantifying Energy and Productivity Benefits of Smart Manufacturing Technologies	2019	C	I4.0, Energy
NS5	[44]	Data Resources to Create Digital Twins	2020	C	DTs, Data
NS6	[45]	On the Engineering of IoT-Intensive Digital Twin Software Systems	2020	C	DTs, IoT
NS7	[46]	Digital Twin Based Software Design in eHealth—A New Development Approach for Health/Medical Software Products	2020	C	DTs eHealth
NS8	[47]	A Taxonomy of Digital Twins	2020	C	DTs
NS9	[48]	Defining infrastructure requirements for the creation of Digital Twins	2020	T	DTs Requirements
NS10	[49]	Developing a Framework for Scoping Digital Twins in the Process Manufacturing Industry	2020	J	DTs Framework
NS11	[50]	A nova agenda da grande indústria: uma análise da indústria 4.0 com base em documentos e materiais de divulgação do projeto alemão plattform industrie 4.0	2020	T	I4.0
NS12	[51]	Modernização de Arquiteturas de Sistemas para suporte à Transformação Digital	2020	C	Systems
NS13	[52]	Toward to Operationalization of Socio-Technical Ontology Engineering Methodology	2020	J	Ontology
NS14	[53]	Applications of ontologies in requirements engineering: a systematic review of the literature	2016	J	Ontologies
NS15	[54]	Integrating Virtual Reality and Digital Twin in Circular Economy Practices: A Laboratory Application Case	2020	C	DTs, CE
NS16	[55]	Software requirements specification analysis using natural language processing technique	2013	J	Requirements

3. Literature Review and Results

This section presents the papers relevant to (i) the identification of DT sustainability requirements, and (ii) the identification of the relationship between DTs and product design. These two dimensions allowed the mapping of the answers we consider in this paper. Papers were sorted along the lines of these two dimensions being selected by the dominant work in each research.

3.1. Digital Twins and Sustainability Requirements

In our set of selected papers, there are 4 SLR papers and 4 non-SLR papers mainly related to DTs and sustainability.

Pokhrel, Katta, and Palacios [3] (S3) study the definition of DT and "state-of-the-art" on the development of DT, including reported work on the usability of a DT for cybersecurity using the SRL methodology. Regarding incident prediction, the cases of the reported use of DTs are: intrusion detection; anomaly detection; monitoring (remote and on-site); virtual commissioning; autonomy; predictive analytics; documentation; and communication. Security is a major dimension of sustainability; for example, if equipment is dangerous its daily usage is probably impossible; their paper is an example of relevant SLR application in the field.

Rosa et al. [10] (S5) assess the relationships between CE and I4.0 using the SLR methodology. They stress the hybrid categories like Circular I4.0 and Digital CE, but move forward to the identification of the main benefits of integrating CE and I4.0, such as production technologies, financial performance, market expansion, supply chain management, product life-cycle management, workforce empowerment, and business models.

Ejsmont, Gladysz, and Kluczek [33] (S6) use a bibliometric literature review to evaluate the impact of Industry 4.0 on sustainability. They find that authors who deal with CE usually also study sustainable supply chains; nevertheless, I4.0 concepts such as sustainability, big data, smart manufacturing, IoT, sustainable development, digital transformation, and industrial IoT are frequently addressed. Cyber-physical systems, sustainable manufacturing, the smart factory, and digitalization are also popular concepts. Possibly, the main conclusion of their paper is that the positive sustainability outcome of these technologies is not guaranteed, and so, success requires supportive measures and specific policies to ensure the competitiveness of local actors.

Rajput and Singh [34] (S7) present an Industry 4.0 model for CE and cleaner production. Their model is built with mixed-integer linear programming (MILP) to optimize product machine allocation, e.g., optimizing the trade-off between energy consumption and machine processing cost. In this model, sensors are also deployed to capture real-time information in the Industry 4.0 facility.

Iñigo [41] (NS2) stresses the complexity of the new I4.0 and the new tools associated with it, for example, 3D printing, to allow the optimization of manufacturing.

Because there is no production without energy (meaning electricity at the production factory level, non-renewable sources of energy still being the paramount contribution to electricity production), the next three papers (i.e., NS1, NS3, and NS4) look to the general requirement of the "responsible use of energy".

Nimbalkar et al. [8] (NS1) and Supekar et al. [43] (NS4) present a framework for quantifying the energy and productivity benefits of smart manufacturing technologies. Breweries are the example used to demonstrate this framework and the implementation of smart manufacturing technologies. To determine the feasibility of a set of smart manufacturing interventions, the framework uses the cost of conserving energy (CCE) as a complementary measure. The quantification and analysis of energy productivity is its focus, and a strategic analysis framework to estimate cost-effective improvements in energy efficiency and productivity, using smart manufacturing, has been developed.

Saad, Faddel, and Mohammed [42] (NS3) study the effective and efficient implementation and design of DTs for energy cyber-physical systems. With the emergence of distributed energy resources (DERs), with communication and control complexities, it is fundamental to guarantee an efficient platform that can digest all the incoming data and ensure the reliable operation of the power system. To build this support technology, two DT types are introduced: one to cover the high-bandwidth applications, and another to the low-bandwidth applications that need centric oversight decision-making. The validation and test of this approach were performed using Amazon Web Services (AWS) as a cloud host that incorporates physical and data models, and additionally can receive live measurements.

3.2. Digital Twins and Product Design

The following papers are important to establish a bridge between DTs and product design. These papers introduce and analyze aspects that should be considered in real scenarios when designing or building DTs. In our set of selected papers, there are 9 SLR papers and 12 non-SLR papers mainly related to DTs and product design:

Jones et al. [1] (S1) try to characterize the DT concept using an SLR. The authors acknowledge that there are a variety of definitions employed across industry and academia. They identified 13 characteristics of DTs to clarify the definition, namely: physical entity/twin; virtual entity/twin; physical environment; virtual environment; state; realization; metrology; twinning; twinning rate; physical-to-virtual connection/twinning; virtual-to-physical connection/twinning; physical processes; and virtual processes and a complete framework and operation process of the DT.

Liu et al. [2] (S2) provide a literature review on DTs based on concepts, technologies, and industrial applications. They evaluate the current state of the art, discuss the concept of the DT, and analyze certain key enabling technologies of DTs. Additionally, they discuss fifteen industrial applications with their respective life cycles, and also present valuable observations and future work recommendations for DT research.

Josifovska, Yigitbas, and Engels [4] (S4) develop a reference framework for DTs within cyber-physical systems (CPSs). The authors define CPSs as system representations that integrate physical units and processes with computational entities over the internet, allowing ubiquitous access to information and services. The framework establishes a relationship between the 5-level CPS architecture and the DT framework, to answer open questions and challenges on how to design and realize CPSs.

Barth et al. [35] (S8) systematize DTs, creating an ontological and conceptual framework. Furthermore, these authors try to answer three research questions: (i) Which dimensions are used to classify and structure DTs in academic literature? (ii) What are the fundamental differences or specifications within these dimensions? and (iii) How do these different specifications relate to each other?

Rub and Bahemia [36] (S9) try to understand the current reality of smart factory implementation using an SLR. They identify a research gap related to the make-or-buy decision around DTs and other core components of the smart factory. This is significant because it is assumed that the smart factory leads to the creation of value, but that creation depends upon the way the factory is implemented—whether the implementation project is executed in-house or using an external supplier.

Dave et al. [37] (S10) discuss the new possible reality of smart cities where concepts such as IoT, big data, AI, robotics, and DTs are paramount, and stress the role of the latter. Smart cities, manufacturing, and healthcare are considered the main fields of application for DTs. An example of the application of DT is their usage in traffic management systems, using traffic cameras that are merely recording, but their recordings can be used to create traffic management models to reduce traffic congestion; they certainly provide more data to export and update a road network with real-time decisions. The authors conclude that there is a need for demonstration sites to test the new technologies with real data, and a need for extensive professional panels of experts in diverse research fields, for example, urban development, IT, transportation, and environmental policies.Polini and Corrado [38] (S11) present an example of DT for the stone-sawing process. The authors describe the DT, but have concerns regarding the accuracy of the equipment and its efficiency and efficacy.

Strmecki et al. [39] (S12) use an SLR to study the possible application of ontologies in automatic programming. Ontologies, which are typically considered as a technique or an artifact used in one or more software life-cycle phases, may be used to help achieve the goal of finding higher abstraction levels, and ways to reuse software to increase its productivity and quality, within the discipline of software engineering.

Sjarov et al. [40] (S3) study the DT concept in the industry in a systematic way. The authors acknowledge a significant growth in the number of scientific studies since 2015 (industry-related publications per year carrying "Digital Twin" in their title). Studies show a variety of applications of DTs ranging from products and processes to whole production systems. Explicit definitions were found to be partly conflicting, and similar notions like "Product Avatar" and "Digital Shadow" are also identified. Their paper extends the theoretical foundation, setting a basis for future, improved DT modeling.

Schweiger, Barth, and Meierhofer [44] (NS5) focus their work on the data resources needed to create DTs. DTs are considered one of the key technologies for organizations moving from producing goods to offering services. The main hypothesis is that a large part of the new data produced, or resources, are already generated during the beginning of life (BOL) phase of the product life cycle, but are not used in the middle of life (MOL) phase. The new framework allows a better understanding of how to use data resources from BOL phases in MOL phases, and permits the creation of an ontology of product data, making the creation and maintenance of DTs easier.

Rivera et al. [45] (NS6) look closely at the engineering of IoT-intensive DT software systems. The authors assume the real DT to be a product that is equipped with several sensors or computing devices that generate, consume, and transfer data for different purposes. Due to this reality, they consider DTs, to a large extent, IoT-intensive systems.

Lutze [46] (NS7) studies the DT-based software design in eHealth as a new development approach for health/medical software products. The author's DT concept builds on (i) a personal digital twin as a Gemini of the patient, (ii) a group digital twin modeling the designated user group of the software, and (iii) a system digital twin for the software product itself. Agile development techniques in comparison to the V-model-based classic software development are considered as offering better support possibilities.

Valk et al. [47] (NS8) present a taxonomy of DTs using an SLR. To accomplish this task several dimensions of DTs are pointed out: data link, purpose, conceptual elements, accuracy, interface, synchronization, data input, and creation time.

Jay [48] (NS9) identifies the infrastructure requirements for the creation of DTs. In addition, the author assumes that after 2015, simulation is a core functionality of systems using seamless assistance along their entire life cycle, i.e., supporting operation and service with a direct link to operation data.

Perno and Hvam [49] (NS10) investigate the processes of the manufacturing industry and develop a framework for scoping DTs in that context. Due to the novelty of the concept and the broad range of technologies upon which it is built, the process of scoping Digital Twin projects can prove to be daunting for process-manufacturing companies.

Arias [50] (NS11) entitles the new project "Plattform Industrie 4.0", and, after briefly describing its technologies, considers its implications.

Leon and Horita [51] (NS12) try to overcome two challenges: (i) how to decentralize existing legacy systems to provide a technology solution that meets the new needs of users in this more digital society; and (ii) how to create a systems architecture that addresses the characteristics inherent in digital transformation.

Sensuse et al. [52] (NS13) use SLR to identify qualitative research in ontology engineering. The main purpose of these authors is an operationalization of socio-technical ontology engineering methodology. This methodology consists of five main phases, namely: (i) planning, (ii) analysis, (iii) design, (iv) implementation, and (v) evaluation.

Dermeval [53] (NS14) uses SLR to identify the applications of ontologies in RE (requirements engineering). The main findings of this research are that: (i) there is empirical evidence of the benefits of using ontologies in RE activities, especially for reducing ambiguity, inconsistency, and the incompleteness of requirements; (ii) the RE process is usually only partially addressed, for example, only considering functional requirements; (iii) ontologies support a great diversity of RE modeling styles; (iv) several studies describe the use/development of tools to support different types of ontology-driven RE approaches; (v) about half of the studies followed W3C recommendations on ontology-related languages; (vi) a great variety of RE ontologies were identified; nevertheless, none of them has been broadly adopted; and (vii) several promising research opportunities were identified. Other authors also have some valuable inputs to this discussion [3,13,37].

Rocca et al. [54] (NS15) try to put into the same basket VR, DTs and CE practices, and present a laboratory application case: virtually testing waste from electrical and electronic equipment (WEEE) disassembly plant configuration, using a set of dedicated simulation tools. The authors stress the importance of their work due to the increasing awareness of customers toward climate change effects, the high demand instability affecting several industrial sectors, and the fast automation and digitalization of production systems.

Fatwanto [55] (NS16) proposes a software requirements specification analysis using natural-language processing techniques. The author tries to improve the software product production process.

4. Discussion

The study of DTs is recent, mainly after 2015 [33,40], and there is an unclear definition of DT. However, regardless of the growing complexity of their applications, it is agreed that there are several benefits to using DTs, such as optimization of Industry 4.0, and the sustainability of the product design process. In addition, the lack of studies with technical details can be a difficulty when adopting this technology. A closer look into possible environmental sustainability benefits and at the product design level can help a clear understanding of this reality.

The analysis of the available literature allows us to identify several aspects of the relationship between DTs, product design and sustainability, and hence, to answer the original research question, we preliminary discuss the involved sub-questions (SQi):

SQ1: What is the relationship between DTs and product design?

Concerning SQ1, from the set of selected papers we verify that there are primarily two relationships between DTs and product design: (i) DTs are digital models of physical products fed with real-time data, having an important role in understanding real behaviors and needed adaptations, and (ii) tests using DTs are less expensive and easier than building new physical prototypes.

SQ2: What are the environmental sustainability requirements of DTs?

Concerning SQ2, and based on the selected literature, the main environmental classes of sustainability requirements are: (i) control of energy consumption and (ii) use of environmentally friendly materials. When CE at the I4.0 level is considered, there is a clear need for a tradeoff between complexity and energy consumption versus the results of the new technology implementation. Possibly, making DT tests and creating less complex products is a main topic of research.

SQ3: What are the open issues and challenges in future research paths for DTs and sustainability?

Concerning SQ3, at a first look and as already mentioned, the complexity reduction to build and set up DTs is only one future research path. The impact of complexity and energy consumption is paramount when the decision to introduce DTs at the factory level is considered. This has an overall impact on sustainability in its dimensions of security, environmental and financial sustainability. The discussed meta-SLR allowed us to identify several possible open issues and challenges for research.

First, methods and processes to design and implement DTs are needed. Since there are several specific application domains with their characteristics, this should be considered. Although there are detailed descriptions of DTs [38,56], it is unclear how to realize them, especially if there is no previous experience on how to do it.

Second, SLR studies point out gaps such as: (i) perceived benefits have not been identified; (ii) DT across the product life cycle or the DT life cycle is not sufficiently studied (whole life cycle, evolving digital profile, historical data); (iii) DTs have not been created, thus, it is not clear how DTs contribute to reducing cost or improving service or supporting decision making; (iv) technical implementations must be improved and detailed in the context of IoT; (v) the level of fidelity is not evaluated in terms of the number of parameters, their accuracy and levels of abstraction; (vi) data ownership of data stored within the DT must be determined; and (vii) integration between virtual entities must be improved, because better methods are needed for communication [1]. As already mentioned, DTs are a recent technology, and this fact partially explains these gaps or diversified future research paths.

At the same time, there is an attempt to clearly define the DT concept, for example, to classify existing standards such as the "Plattform Industry 4.0", which describes a standardized DT in I4.0 [47]. If there are several approaches and contexts to define DTs, how can we identify the most important gaps in the study of sustainability in DT usage? We start with the assumption that a DT consists of three parts: (i) a physical product, (ii) a virtual product, and (iii) connections and data flowing between them [47]. Then, we assume that at the technical implementation and at a particular level of fidelity, it is possible to identify its main contributions to sustainability.

RQ1: What is the state of the art in the area of sustainability requirements of DT-based systems related to product design?

Finally and trying to answer our original research question RQ1, we must address it carefully. Despite the existing gaps, the literature identifies several sustainability requirements for DT-based systems related to product design, namely: (i) fidelity; (ii) energy control; (iii) complexity control; (iv) identification of environmentally friendly and cost-efficient materials, and (v) easy reproduction of new product designs. Different studies investigate distinct sustainability requirements, and there is no integrated approach to understanding how DTs can create environmental sustainability. An integrated approach would imply additional complexity, for example, a fully rigorous fidelity implies additional time and further energy consumption, and this output might create a tradeoff, leading to fuzzy fidelity. This fuzzy fidelity means that environmental costs continue to be external to the production because their evaluation would also imply further work and costs. Easy reproduction of new product designs might imply a reduction of costs at the production stage, but the costs of the first steps of DT implementation might explain why this requirement is such a demanding one.

5. Conclusions

In this work, it has been possible to identify relevant research work regarding the study of DT-based systems and technologies, using the SLR methodology as the main tool for a meta-analysis on the subject of SLR. Special attention was put on the choice of vocabulary used to perform the research in several databases. Based on that analysis, it was possible to answer the RQ1 as well as the sub-questions SQ1, SQ2, and SQ3.

There are five main concerns to address in the development of a sustainable DT: (i) fidelity; (ii) energy control; (iii) complexity control; (iv) identification of environmentally and cost-efficient materials; and (v) easy reproduction of new product designs. It is also possible to identify areas of research related to DTs, namely: (i) the study of its concept and definition [1,4,35,36,40,47]; (ii) the presentation of examples [38,56]; and (iii) the use of SLR to understand the current research and to identify future research paths [2,10,33,37].

This analysis allows us to identify two main gaps that correspond to two future research paths. The first gap is the absence of a detailed paper explaining exactly what a DT is, with an extensive and rich example that stresses even the hardware characteristics, and how the connections between the physical and digital dimensions can be designed, developed, and maintained. This reality might be explained by the unwillingness of the industry to disclose sensitive information, this gap to be fulfilled by the academy.

Secondly, we identify papers that present sustainability in DT applications, and, additionally, we identify studies that distinguish between several types of sustainability, including environmental sustainability. Furthermore, we identify papers where the connection between CE and the usage of DTs, in the context of Industry 4.0, is clearly stated. However, it was not possible to find a paper that only discusses an SLR of DTs regarding environmental sustainability. Is this merely a vocabulary issue, CE being equal to environmental sustainability? We believe the answer is no, because CE, at the factory level, is still a concept at the laboratory stage, and the complexity implied by its implementation might be paramount. In other words, for the question of whether CE is environmentally sustainable if the management and technical complexity to achieve it is so noteworthy, to evaluate this hypothesis, a future research path might be the study of a scenario where a zero environmental impact product as a control is the design objective, with the help of a DT.

In summary, this paper presents a meta-SLR on DT-based systems that allow us to identify and discuss the main classes of requirements to consider in the development of a sustainable DT, but also allow us to identify gaps and limitations in both research and practice aspects.

Author Contributions: Conceptualization, All; formal analysis, R.C.; funding acquisition, A.R.d.S.; investigation, All; evaluation methodology, R.C.; supervision, A.R.d.S.; writing—original draft, R.C.; writing—review and editing, All. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by Portuguese national funds through FITEC-Programa Interface, with reference CIT "INOV-INESC Inovação-Financiamento Base", and FCT UIDB/50021/2020. The APC was funded by FITEC-Programa Interface, with reference CIT "INOV-INESC Inovação-Financiamento Base".

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: Work supported by funds under FITEC-Programa Interface, CIT INOV- INESC Inovação-Financiamento Base, and FCT UIDB/50021/2020.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

Concept	Acronym
Artificial Intelligence	AI
Circular Economy	CE
Controlled Vocabularies	CV
Digital Twins	DT
Flexible Manufacturing System	FMS
Industry 4.0	I4.0
Information Technology	IT
Internet of Things	IoT
Requirements Engineering	RE
Research Question	RQ
Sub-questions	SQ
Systematic Literature Review	SLR

References

1. Jones, D.; Snider, C.; Nassehi, A.; Yon, J.; Hicks, B. Characterising the Digital Twin: A systematic literature review. *CIRP J. Manuf. Sci. Technol.* **2020**, *29*, 36–52. [CrossRef]
2. Liu, M.; Fang, S.; Donga, H.; Xu, C. Review of digital twin about concepts, technologies, and industrial applications. *J. Manuf. Syst.* **2020**, 346–361. [CrossRef]
3. Pokhrel, A.; Katta, V.; Colomo-Palacios, R. Digital Twin for Cybersecurity Incident Prediction: A Multivocal Literature Review. In Proceedings of the 2020 IEEE/ACM 42nd International Conference on Software Engineering Workshops (ICSEW), Seoul, Korea, 27 June–19 July 2020; pp. 671–678.

4. Josifovska, K.; Yigitbas, E.; Engels, G. Reference Framework for Digital Twins within Cyber-Physical Systems. In Proceedings of the IEEE/ACM 5th International Workshop on Software Engineering for Smart Cyber-Physical Systems (SEsCPS), Montreal, QC, Canada, 28 May 2019; pp. 25–31.
5. Minerva, G.; Lee, M.; Crespi, N. Digital Twin in the IoT Context: A Survey on Technical Features, Scenarios, and Architectural Models. *Proc. IEEE* **2020**, *108*, 1785–1824. [CrossRef]
6. Sepasgozar, S. Digital Twin and Web-Based Virtual Gaming Technologies for Online Education: A Case of Construction Management and Engineering. *Appl. Sci.* **2020**, *10*, 4678. [CrossRef]
7. Verdouw, C.; Tekinerdogan, B.; Beulens, A.; Wolfert, S. Digital twins in farming systems. *Agric. Syst.* **2021**, *189*, 103046. [CrossRef]
8. Nimbalkar, S.; Supekar, S.; Meadows, W.; Wenning, T.; Guo, W.; Cresko, J. *Enhancing Operational Performance and Productivity Benefits by Implementing Smart Manufacturing Technologies in Breweries*; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2019; pp. 93–107.
9. UN. Our Common Future. The Brundtland Report, Oslo. 1987. Available online: https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf (accessed on 2 January 2021).
10. Rosa, P.; Sassanelli, C.; Urbinati, A.; Chiaroni, D.; Terzi, S. Assessing relations between Circular Economy and Industry 4.0: A systematic literature review. *Int. J. Prod. Res.* **2020**, *58*, 1662–1687. [CrossRef]
11. Samadi, E.; Kassou, I. The Relationship between IT and Supply Chain Performance: A Systematic Review and Future Research. *Am. J. Ind. Bus. Manag.* **2007**, *6*, 480–495. [CrossRef]
12. Gereco, A.; Caterino, M.; Fera, M.; Gerbino, S. Digital Twin for Monitoring Ergonomics during Manufacturing Production. *Appl. Sci.* **2020**, *10*, 7758. [CrossRef]
13. Massey, J. The Sumptuary Ecology of Buckminster Fuller's Designs. In *A Keener Perception, Ecocritical Studies in American Art History*; University Alabama Press: Tuscaloosa, AL, USA, 2009; pp. 189–212.
14. Merten, T.; Schäfer, T.; Bürsner, S. Using RE knowledge to assist automatically during requirement specification. In Proceedings of the 2012 Seventh IEEE International Workshop on Requirements Engineering Education and Training (REET), Chicago, IL, USA, 24 September 2012; pp. 9–13.
15. Paech, B.; Moreira, A.; Araujo, J.; Kaiser, P. Towards a Systematic Process for the Elicitation of Sustainability Requirement. In Proceedings of the CEUR Workshop Proceedings, 8th International Workshop on Requirements Engineering for Sustainable Systems, RE4SuSy 2019, Jeju, Korea, 24 September 2019.
16. Maciel, D.; Paiva, A.; Rodrigues da Silva, A. From Requirements to Automated Acceptance Tests of Interactive Apps: An Integrated Model-based Testing Approach. In Proceedings of the ENASE'2019, SCITEPRESS, Heraklion, Crete, Greece, 3–5 May 2019.
17. Paiva, A.; Maciel, D.; Silva, A. From Requirements to Automated Acceptance Tests with the RSL Language. In *Evaluation of Novel Approaches to Software Engineering (ENASE 2019)*; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2020; Volume 1172.
18. Caramujo, J.; Silva, A.; Monfared, S.; Ribeiro, A.; Calado, P.; Breaux, T. RSL-IL4Privacy: A Domain-Specific Language for the Specification of Privacy-Aware Requirements. In *Requirements Engineering*; Springer: Cham, Switzerland, 2019; Volume 24. [CrossRef]
19. Silva, A. Rigorous Specification of Use Cases with the RSL Language. In Proceedings of the International Conference on Information Systems Development'2019, AIS, Toulon, France, 28–30 August 2019.
20. Carvalho, R.; Silva, A. Discussion Towards a Library of Software Sustainability Requirements. In Proceedings of the 9th International Workshop on Requirements Engineering for Sustainable Systems (RE4SuSy) at RE'2020, Online, 20–21 September 2020.
21. Gonçalves, L.; Silva, A. Towards a Catalogue of Reusable Security Requirements, Risks and Vulnerabilities. In Proceedings of the International Conference on Information Systems Development'2018, AIS, Lund University, Sweden, 22–24 August 2018.
22. Silva, A. Model-Driven Engineering: A Survey Supported by a Unified Conceptual Model. In *Computer Languages, Systems and Structures, 43(C)*; Elsevier: New York, NY, USA, 2015; pp. 139–155. [CrossRef]
23. Silva, A.; Sequeira, C. Towards a Library of Usability Requirements. In Proceedings of the ACM SAC'2020 Conference, ACM, Brno Czech Republic, 30 March–3 April 2020.
24. Kitchenham, B.; Brereton, P.; Budgen, D.; Turner, M.; Bailey, J.; Linkman, S. Systematic literature reviews in software engineering—A systematic literature review. *Inf. Softw. Technol.* **2009**, *51*, 7–15. [CrossRef]
25. Escallón, R.; Aldea, A. On Enterprise Architecture Patterns: A Systematic Literature Review. In Proceedings of the 22nd International Conference on Enterprise Information Systems (ICEIS 2020), Online, 5–7 May 2020; pp. 666–678.
26. Wolfswinkel, J.; Furtmueller, E.; Wilderom, C. Using Grounded Theory as a Method for Rigorously Reviewing Literature. *Eur. J. Inf. Syst.* **2013**, *22*, 45–55. [CrossRef]
27. Ahmad, A.; Justo, J.; Feng, C.; Khan, A. The Impact of Controlled Vocabularies on Requirements Engineering Activities: A Systematic Mapping Study. *Appl. Sci.* **2020**, *10*, 7749. [CrossRef]
28. Leise, F. Controlled vocabularies: An introduction. *Indexer* **2008**, *26*, 121–126. [CrossRef]
29. Silva, A.; Savić, D. Linguistic Patterns and Linguistic Styles for Requirements Specification: Focus on Data Entities. *Appl. Sci.* **2021**, *11*, 4119. [CrossRef]

30. Polpinij, J.; Ghose, A. An automatic elaborate requirement specification by using hierarchical text classification. In Proceedings of the 2008 International Conference on Computer Science and Software Engineering, Hubei, China, 12–14 December 2008; pp. 706–709.
31. Pizard, S.; Vallespir, D. Towards a controlled vocabulary on software engineering education. *Eur. J. Eng. Educ.* **2017**, *42*, 927–943. [CrossRef]
32. NISO. ANSI/NISO Z39.19-2005 (2010): Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies, NISO. Available online: https://www.niso.org/publications/ansiniso-z3919-2005-r2010 (accessed on 30 July 2019).
33. Ejsmont, K.; Gladysz, B.; Kluczek, A. Impact of Industry 4.0 on Sustainability—Bibliometric Literature Re-view. *Sustainability* **2020**, *12*, 5650. [CrossRef]
34. Iñigo, I. A Review of Industry 4.0 Potential to Accelerate to Accelerate the Transition to a Circular Economy. Master's Thesis, Industrial Engineering, Universidad de Pais Vasco, Escuela de Ingeniería de Bilbao, Bilbao, Spain, 2020; pp. 1–122.
35. Saad, A.; Faddel, S.; Mohammed, O. IoT-Based Digital Twin for Energy Cyber-Physical Systems: Design and Implementation. *Energies* **2020**, *13*, 4762. [CrossRef]
36. Rajput, S.; Singh, S. Industry 4.0 Model for circular economy and cleaner production. *J. Clean. Prod.* **2020**, *277*, 123853. [CrossRef]
37. Supekar, S.; Graziano, D.; Riddle, M.; Nimbalkar, S.; Das, S.; Shehabi, A.; Cresko, A. A Framework for Quantifying Energy and Productivity Benefits of Smart Manufacturing Technologies. In Proceedings of the 26th CIRP Life Cycle Engineering (LCE) Conference, Purdue University, West Lafayette, IN, USA, 7–9 May 2019; pp. 699–704.
38. Barth, L.; Ehrat, M.; Fuchs, R.; Haarmann, J. Systematization of Digital Twins: Ontology and Conceptual Framework. In Proceedings of the ICISS 2020, Cambridge, UK, 19–22 March 2020; pp. 13–23.
39. Schweiger, L.; Barth, L.; Meierhofer, J. Data Resources to Create Digital Twins. In Proceedings of the 2020 7th Swiss Conference on Data Science (SDS), KKL Luzern, Switzerland, 26 June 2020; pp. 55–56.
40. Rivera, L.; Müller, H.; Villegas, N.; Tamura, G.; Jiménez, M. On the Engineering of IoT-Intensive Digital Twin Software Systems. In Proceedings of the 2020 IEEE/ACM 42nd International Conference on Software Engineering Workshops (ICSEW), Seoul, Korea, 27 June–19 July 2020; pp. 631–638.
41. Lutze, R. Digital Twin Based Software Design in eHealth—A New Development Approach for Health/Medical Software Products. In Proceedings of the 2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Cardiff, UK, 15–17 June 2020; pp. 1–9.
42. Valk, H.; Haße, H.; Möller, F.; Arbter, M.; Henning, J.; Otto, B. A Taxonomy of Digital Twins. In Proceedings of the Americas Conference on Information Systems, Online, 10–14 August 2020; pp. 1–10.
43. Jay, M. Defining Infrastructure Requirements for the Creation of Digital Twins. Master's Thesis, School of Innovation, Design and Engineering, RISE Research Institutes of Sweden, Malarden University Sweden, Västerås, Sweden, 2020; pp. 1–47.
44. Perno, M.; Hvam, L. Developing a Framework for Scoping Digital Twins in the Process Manufacturing Industry. 2020, pp. 1–12. Available online: Media.sps2020.se (accessed on 2 January 2021).
45. Rüb, J.; Bahemia, H. A Review of the Literature on Smart Factory Implementation. In Proceedings of the 2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Valbonne Sophia-Antipolis, France, 17–19 June 2019; pp. 1–9.
46. Arias, A. A Nova Agenda da Grande Indústria: Uma Análise da Indústria 4.0 com Base em Documentos e Materiais de Divulgação do Projeto Alemão Plattform Industrie 4.0. Ph.D. Thesis, Universidade Tecnológica Federal do Paraná Programa de Pós-Graduação em Tecnologia e Sociedade, Curitiba, Brazil, 2020; pp. 1–163.
47. Leon, P.; Horita, F. Modernização de Arquiteturas de Sistemas para suporte à Transformação Digital. In Proceedings of the Anais Estendidos do XVI Simpósio Brasileiro de Sistemas de Informação, SBC, Online, 3–6 November 2020; pp. 61–66.
48. Sensusea, D.; Sucahyoa, Y.; Silalahia, M.; Wulandaria, I.; Akmaliaha, F.; Noprisson, H. Toward to Opera-tionalization of Socio-Technical Ontology Engineering Methodology. In Proceedings of the 2017 5th International Conference on Cyber and IT Service Management (CITSM), Denpasar, Indonesia, 8–10 August 2020; pp. 1–7.
49. Dermeval, D.; Vilela, J.; Bittencourt, I.; Castro, J.; Isotani, S.; Brito, P.; Silva, A. Applications of ontologies in re-quirements engineering: A systematic review of the literature. *Requir. Eng.* **2016**, *21*, 405–437. [CrossRef]
50. Dave, R.; Dave, S.; Thakkar, H. Digital Twins: Current problems in Smart City and Recommendations for future technology. *Int. Res. J. Eng. Technol. (IRJET)* **2020**, *7*, 1–11.
51. Polini, W.; Corrado, A. Digital twin of stone sawing processes. *Int. J. Adv. Manuf. Technol.* **2020**, *112*, 121–131. [CrossRef]
52. Strmečki, D.; Magdalenić, I.; Radosević, D. A Systematic Literature Review on the Application of Ontologies in Automatic Programming. *Int. J. Softw. Eng. Knowl. Eng.* **2018**, *28*, 559–591. [CrossRef]
53. Rocca, R.; Rosa, P.; Sassanelli, C.; Fumagalli, L.; Terzi, S. Integrating Virtual Reality and Digital Twin in Circular Economy Practices: A Laboratory Application Case. *Sustainability* **2019**, *12*, 2286. [CrossRef]
54. Sjarov, M.; Lechler, T.; Fuchs, J.; Brossog, M.; Selmaier, A.; Faltus, F.; Donhauser, T.; Franke, J. The Digital Twin Concept in Industry—A Review and Systematization. *IEEE Explor.* **2020**, *1*, 1789–1796.
55. Fatwanto, A. Software requirements specification analysis using natural language processing technique. In Proceedings of the 2013 International Conference on QiR, Yogyakarta, Indonesia, 25–28 June 2013; pp. 105–110.
56. Pang, T.; Restrepo, J.; Cheng, C.-T.; Yasin, A.; Lim, H.; Miletic, M. Developing a Digital Twin and Digital Thread Framework for an 'Industry 4.0' Shipyard. *Appl. Sci.* **2021**, *11*, 1097. [CrossRef]

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

Applied Sciences Editorial Office
E-mail: applsci@mdpi.com
www.mdpi.com/journal/applsci

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland

Tel: +41 61 683 77 34
Fax: +41 61 302 89 18

www.mdpi.com

ISBN 978-3-0365-1799-5