Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 116 (2023) 143–148

www.elsevier.com/locate/procedia

30th CIRP Life Cycle Engineering Conference

Sustainability Digital Twin: a tool for the manufacturing industry

Maria F. Davila R.*[a], Fenja Schwark[a], Lisa Dawel[a], Dr. Alexandra Pehlken[a]

[a] *OFFIS Institute for Information Technology, Escherweg 2, Oldenburg 26121, Germany*

* Corresponding author. Tel.: +49-441-9722-744 ; fax: +49-441-9722-278. *E-mail address:* maria.davila@offis.de

Abstract

The use of Digital Twins in the manufacturing industry has significantly increased in the past years, given its potential to contribute to solutions for the current environmental and digitalization challenges. To address the needs of the industry and achieve a higher level of sustainability a robust tool, such as a Digital Twin, must be used because the physical processes are highly complex, with constantly changing variables. On the other hand, current goals of the manufacturing sector, such as circular economy and zero-emissions, require data sharing in the supply chain. This paper presents the development of a framework for the assessment of energy efficiency and sustainability of production processes, with respect to the measured consumption of the different resources. In addition, a concept for privacy preserving data sharing within the manufacturing supply chain is presented. Two real case studies from the big component production industry located in northern Germany are discussed in this paper. As result, the framework is presented and the case study successfully identifies significant energy efficiency and sustainability optimization possibilities, achieved with very simple modeling.

Keywords: Digital Twin; Sustainability; Industry 4.0; Privacy-preserving data sharing

1. Introduction

The cost of energy has become significant within production costs and there are supply disruption risks, which results in the urgent need of monitoring and, if possible, optimizing of energy consumption in the manufacturing sector. Additionally, many companies have set environmental goals such as *zero emission* or *circular economy* production. Digital Twins in the manufacturing industry have significant potential to contribute to solutions for the environmental and digitalization challenges. This potential is only possible because the Industry 4.0 comes with higher inter-connectivity and automation in the production process. There is no common understanding of Digital Twins in the literature because it slightly varies in each area of application, nevertheless, for the purposes of the presented study, a Digital Twin is defined as a software system, which uses models and services to represent and manipulate a physical system, with a specific goal in focus. The focus in this case is sustainability.

A great portion of the published work sets electricity consumption as optimization parameter, nevertheless the energy crisis faced today is only an example of the trend followed by other resources such as water, and critical raw materials. Also, a great portion of the energy consumption in production plants is not electricity, but compressed air, gas or heat. This demonstrates the need for more robust tools, which are able to take into account many critical and constantly-changing variables such as profit, electricity, water, gas and raw materials.

The presented work is part of a bigger study in the development of a robust tool in the form of a sustainability Digital Twin. There is a research gap in the manufacturing sector, where the optimization goal of the Digital Twin is sustainability. Companies can't easily implement a Digital Twin because there are no frameworks, which are easy to deploy and generic. Establishing consumption or emission standards is non-trivial, since manufacturing processes in different companies are completely heterogeneous. Therefore, the assessment of energy efficiency and sustainability has to be done with respect to the process itself, by measuring different variables and modeling their behavior, and the result must be information that enables better decision-making within the company. Additionally, zero emissions and circular economy does not only depend on what happens within one company, but it affects the whole supply chain. As a consequence, data privacy concerns become one of the main barriers of such a Digital Twin framework because data sharing becomes necessary.

10.1016/j.procir.2023.02.025

The architecture of the Digital Twin has different levels: connection and communication, data-to-information conversion, cognition or analysis and finally visualization and control [1]. Presenting the whole structure would be out of the scope, hence the central point of this paper are the data-to-information and the analytical levels of the Digital Twin architecture.

To summarize, the present study makes the following contributions. (1) Development of a framework for the assessment of sustainability of production processes, with respect to the measured consumption of the different resources. This includes process modeling, in order to measure the relevant variables and a base model for the assessment. (2) A concept for privacy-preserving data sharing within the manufacturing supply chain initially done with ordinal encoding.

Section 2 presents the relevant related work and theory, to better describe the research gap addressed. Section 3 presents the methods developed and section 4 the results for a proof-of-concept. Finally, section 5 discusses the framework, and concludes including the future work, which will be carried out within the project.

2. Background and Related Work

2.1. Digital Twins in the Manufacturing Sector

Dalibor et al. [2] presented a systematic mapping of 356 Digital Twins publications to characterize their structure in the different domains. 20 different application domains are presented, but 70% of the studied work corresponds to the manufacturing industry. They are used for monitoring and controlling either before production, to improve the design process, or during product lifetime. Kritzinger et al. [3] found that 49% of the work sets production planning and control (PPC) as focus, with production as optimization parameter. Biesinger et al. [4] proposed the creation of a Digital Twin of a production system of the automotive industry to optimize planning cycle times. Liu et al. [5] developed important concepts for the different communication protocols of a Digital Twin in production, such as OPC UA and MTConnect. Lee et al. [6] developed a reference architecture based on deep learning and 5C-CPS ([1]) to facilitate the transformation towards smart manufacturing and Industry 4.0.

2.2. Digital Twins for Energy Efficiency and Sustainability

He et al. [7] created a review of Digital Twin-based sustainable intelligent manufacturing studies. They present an extensive literature review where, even though the title might be misinterpreted, there are no examples of *sustainability* Digital Twins. Instead, the presented studies refer to Digital Twins in the manufacturing sector, whose objectives (for example design [8], production control improvement [9], or equipment status monitoring [10]) could in parallel improve the sustainability of the manufacturing process. This does not mean that sustainability is the optimization variable in those studies, which is the gap addressed in this paper.

Hasselbring et al. [11] first developed Titan, which is a robust control center, integrating and analyzing big data in industrial manufacturing. Afterwards, Henning et al. [12] use the Titan architecture to monitor and analyze the power consumption for two pilot cases. The Digital Twin developed in our project aims to perform similar analyses, but in our case we include other types of energy different from electricity (i.e. gas, heat, compressed air) and relevant resources for sustainability, such as water and raw materials.

2.3. Privacy-Preserving Data Sharing

Privacy-preserving data sharing has been until recently a topic mainly for health data, given that the health sector required extensive data sharing between different actors, such as hospitals, laboratories and pharmacies, and also dealt with very sensible patient information. With the increasing awareness about the relevance of privacy and privacy regulations becoming more strict (for example GDPR [13]), the readiness to share sensitive information has decreased significantly. Simultaneously, the need for data sharing in the manufacturing sector has increased significantly. Concepts like circular economy or zero-emissions production are impossible to achieve without data sharing in the supply chain.

For this reason, a requirement for our sustainability Digital Twin is that the output must take into account the privacy concerns of the different actors. Therefore, after the data collection, it should be pre-processed in a way that does not reveal protected information about the production process or the machines used in it.

There are many approaches to achieve privacy-preserving data sharing, for example encryption [14], anonymization [15], perturbation [16], encoding or more complex data synthesis (creating data synthetically). Choosing the appropriate technique depends mainly on two aspects: (1) how is the data going to be used downstream, and (2) what level of privacy, or disclosure risk desired. In this work, categorical data is encoded, and numerical data is normalized. When dealing with text or categorical data with a limited number of possible values, encoding is a practical method, which avoids disclosing the information behind the categories. In our case study, an example for categorical data are the states of the machine. Revealing the information of the states of a machine makes it possible to reverse-engineer the operating principles of such machine. Future work includes choosing a more complex privacy-preserving method when more sensitive data sources are included.

3. Methodology

The goal of the project is assessing energy efficiency and sustainability of the manufacturing process. As mentioned before, there is currently no unified definition of a Digital Twin in the literature. Therefore, figure shows its architecture per our definition of the concept. The scope of this paper are the levels 2 and 3 of the architecture: the pre-procesing of the data and the analytics.

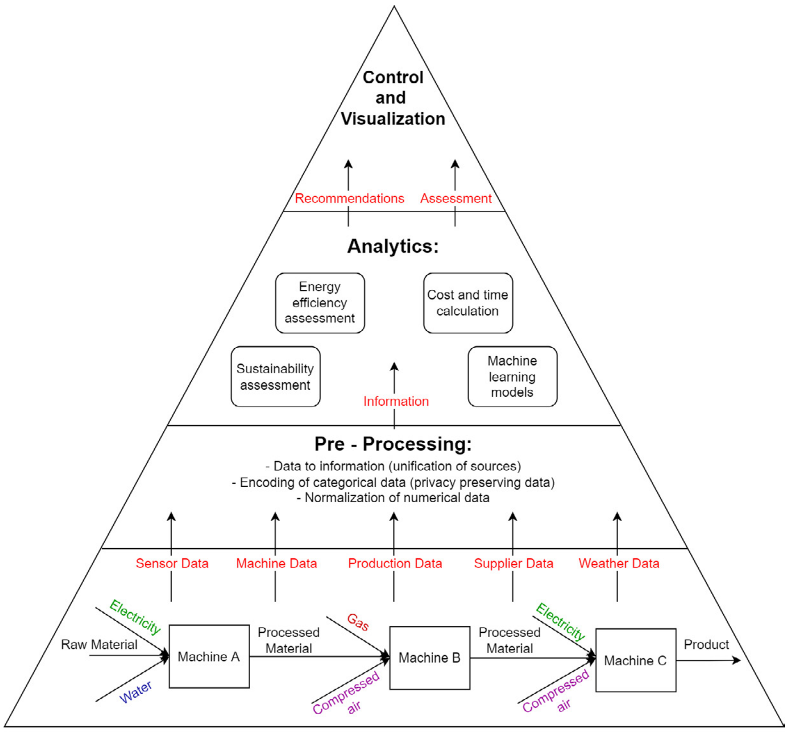

Fig. 1. Architecture of the proposed sustainability Digital Twin

The following explains in detail each of the levels in figure 1. Level 1 is called connection and communication. The development of the Digital Twin starts by creating a process model of the manufacturing steps, as shown in figure 2. The first step is defining the system boundaries, starting with the raw materials and ending with the finished product. Afterwards, all the required **inputs** and **outputs** of each of the machines involved in the different production processes are determined. The different inputs can be electricity, heat, gas, compressed air, or materials, within others. The output are typically the processed materials, intermediate products and finally the product. All the relevant measurement points within the process are determined, in order to install sensors required for the measurement of the inputs. As a result of the process modeling, the flow of resources into and out of the system boundaries can be measured and monitored in real-time.

Level 2 in the development of the Digital Twin, broadly put, is gathering the data measurements from the points of interest configuring the communication between machines and to the other levels of the architecture. These measurements provide many types of different data sources. They include sensor, machine, production, supplier and weather data (relevant for models with seasonal variability and for temperature corrections in compressed air analysis). The sources can have different data types: tabular (numerical, categorical and binary), relational tables, time series data, among others. Pre-processing is done, in order to combine different types of sources. Additionally, categorical data is encoded and numerical data is normalized. The chosen encoding technique is called Ordinal Encoder [17]. In the case of ordinal encoders, the algorithm assumes that two nearby values are more similar to each other than to distant values, which is the case for the operation of a machine, and the reason why we chose it for the machine states [18]. Similarly, min-max normalization is performed to the numerical data. In this proof-of-concept, only sensor, machine and production data is used, therefore, there are no further privacy methods used.

Level 3 of the Digital Twin is the cognitive or analytical level. Here, the information is used to perform energy efficiency and sustainability assessments and the costs and duration per product are calculated. Figure 3 is a diagram showing a practical example. Each of the lanes represents a possible production process for the same product, with the involved machines and their inputs. The right column shows the results of the assessment. The assessment outputs the amount of manpower (relevant for costs and planning), water, electricity, gas, compressed air and heat necessary for each process. Production process 1 is the standard process, since it represents the lowest cost of production. Nevertheless, it represents a high risk because the machines involved require many different types of resources. Production process 2 is the scenario when there is a gas supply disruption, hence machine B can no longer be operated. Therefore, the product is then produced using machines D and E. This means that the process lasts longer and is more expensive, but it does not consume gas. Nevertheless, process 2 has the best sustainability and energy efficiency rating, since the heat bought is generated in a more environmentally friendly way and less energy conversion is needed, hence less losses. Production process 3 is the example of when there is a water shortage and the company is forced to skip a production step and buy the preprocessed materials. In this case, production time is shorter, but costs are higher. Also, sustainability decreases because there are more transportation and supply chain impacts.

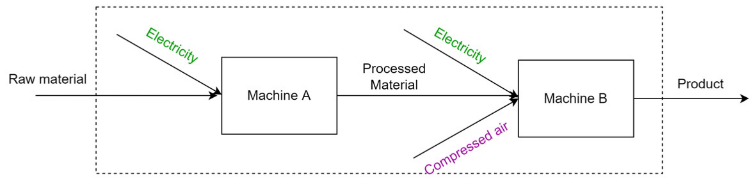

Fig. 2. Process model for the simplified case study with two machines

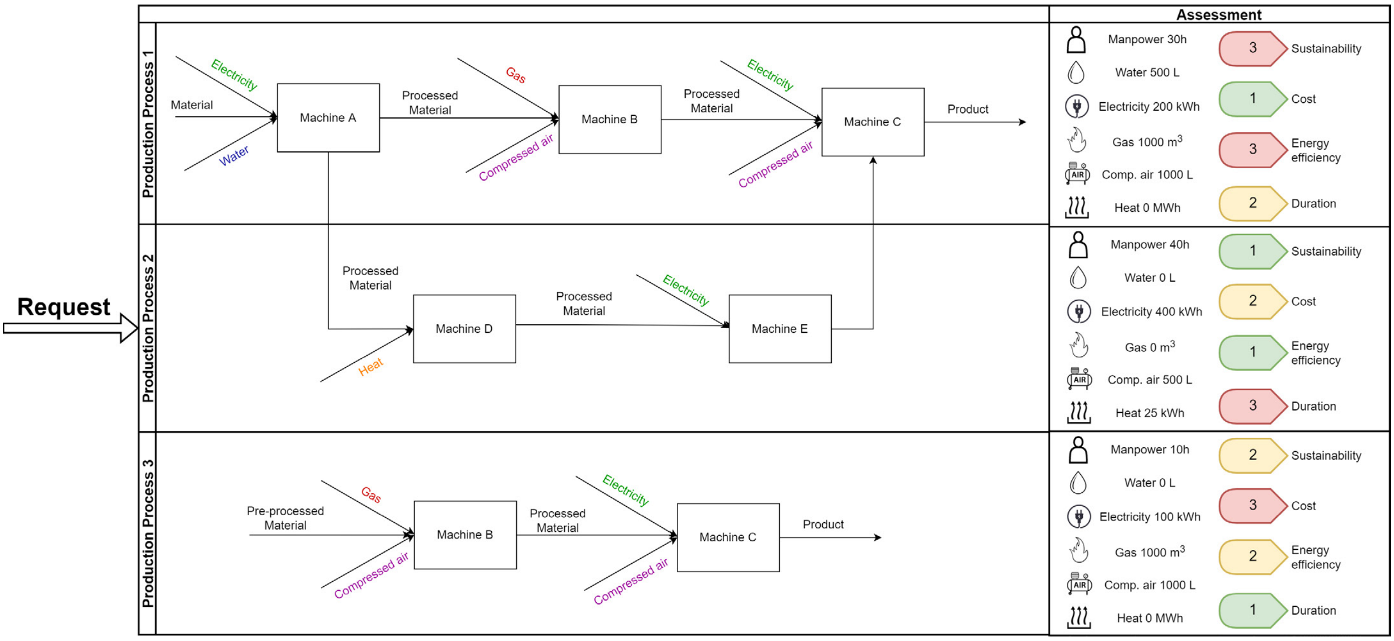

Fig. 3. Example of the assessment performed in the analytical level of the sustainability Digital Twin

Quantifying the costs per product becomes a complex task under such ever-changing scenarios. The proposed Digital Twin framework establishes the appropriate measurement points to perform these calculations, which is optimal because it supports decision-making. Additionally, an assessment is performed, where the current process is compared to historical data and the efficiency is classified into low, average and high. The process is compared with itself and helps identifying anomalies or failures, to optimize consumption. This, combined with the use of resources like water and raw materials, provides an assessment of how sustainable the process is. Emissions can be calculated by knowing how much energy is consumed and the source where it comes from.

Level 4 of the Digital Twin is the visualization of the results of the assessment and, after future work, real-time optimization of the production plan. Visualizing the different calculation of resources (figure 3) has supports the decision-making, to choose between the different production processes, taking into account many different and constantly changing variables. This allows to monitor and optimize the production process. The steps of the sustainability Digital Twin are summarized as:

1. Modeling the manufacturing process steps and defining the input of resources and measurement points.
2. Data transfer - not included in this paper
3. Data pre-processing from all the different data sources and types.
4. Privacy-preserving by encoding categorical machine, sensor and production data and normalizing numerical data types.
5. Performing the energy efficiency and sustainability assessment to support decision-making within different production processes and quantify the impact per product.

3.1. Case Study

The results presented come from a real case study from the big component production industry located in northern Germany for two different companies. The process presented is simplified as it serves as proof-of-concept and can be considered as work in progress. The measurements were done in 2022. For the proof-of-concept, two machines within the process were selected: a riveting machine and a salt-bath furnace. These two machines were chosen as a starting point because they represent the highest energy consumption within the process, also the most time consuming and finally because they have more than one type of resource as input. This and the system boundaries are shown in figure 2. The complete production process includes 10 different machines, for which 8 measurement points need to be installed.

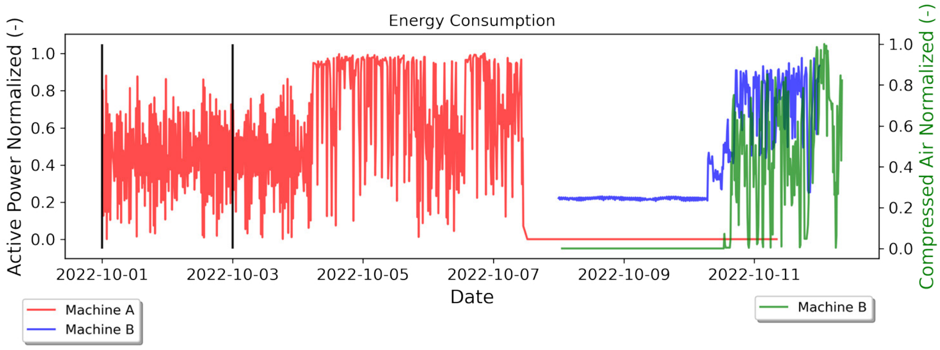

Fig. 4. Energy consumption of machine A and machine B. In red, electric energy consumption of machine A. In blue, electric energy consumption of machine B. In green, compressed air consumption of machine B. Vertical black lines mark the beginning and end of a weekend.

4. Results

4.1. Data to Information

Data from machine A is time series type data, measured every 15 minutes from 50 different sensors (temperature, rotating velocity, power, vibration, etc.). Data from machine B is tabular data captured every time there is a change in the state of the machine (for example, from parking to movement), hence it is not continuous but it has a higher resolution of 1 second. Machine B provide electrical power and compressed air flow measurements. Both sources are pre-processed into continuous time series, with equivalent sampling.

4.2. Privacy-Preserving Conversion Process

The resulting tables from the previous step are processed into privacy-preserving data. For this, all categorical data is encoded using an ordinal encoder as explained in section 2 and the numerical data is normalized using min-max normalization. This avoids revealing sensitive information about the machines and allows to perform the same analysis, since all comparisons are percentual and not absolute values, which also makes more sense for comparison with other production processes.

4.3. Energy Efficiency and Sustainability Assessment

Figure 4 shows the energy consumption of both machines, with respect to time. Machine A requires only electricity and machine B electricity and compressed air. Compressed air is a form of energy storage, therefore it is also considered within the assessment as energy consumption. One can either measure the electricity consumption of the compressor producing the air, or measure the air flow of the machine and convert it using the nominal power of the compressor. This information is then used to integrate and calculate the total energy consumption of each of the machines per product. In machine B, the energy consumption of the compressor represents approximately 18% of the total consumption. This is one clear example of the impact it has, when only electrical energy is considered.

This information allows the planners to identify energy saving possibilities, for example by identifying high energy consumption times, when the machines are not being productive. An example of this can also be seen in figure 4, where the time between the black vertical lines is a weekend, during which there is no production, however, machine A still consumes approximately 40% of the energy it needs when it is being productive.

Better modeling is achieved once production data is included in the analysis. For example, the control system of machine B provides information of the state of the machine (standby, moving, operating, failure) for control purposes. This data is used in combination with the energy consumption data. The different states of the machine were encoded using an ordinal encoder and the result is shown in figure 5. The combination of the two sources result in the correlation between energy consumption and state of the machine. This allows to better quantify the energy used productively, as a measure of the efficiency of the process. For example, in figure 5, code 0 represents the state where the machine is in standby and one can see that the standby energy consumption is close to 0. Code 17 represents the state where the machine is producing and all other states are planned and un-planned stops. During most of those stops, the power consumption should be significantly reduced because it represents energy being wasted. This type of analysis is very helpful, in order to increase the process's efficiency.

With the above, the model is already able to output the costs, energy consumption and duration of the production. This information is being gathered, in order to create a historical basis. Once enough data is available, models as described in section 3 will be implemented, in order to provide a comparison of the energy efficiency and sustainability of the real-time data vs. historical data.

4.4. Visualization

The models shown before are all visualized using Grafana, an open source interactive visualization web application. Custom-made platforms could be adapted to the results of the models.

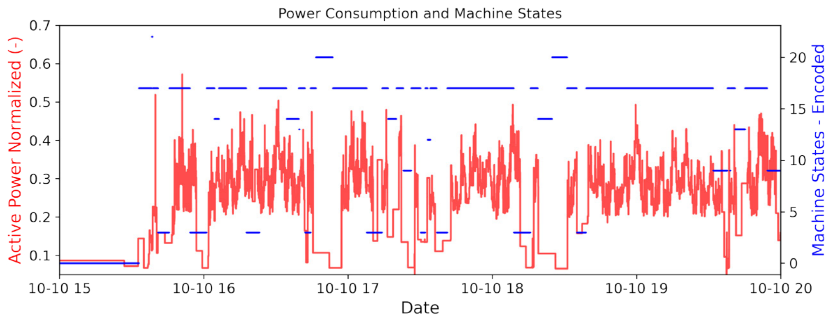

Fig. 5. Energy consumption of machine B correlated with the states of the machine

5. Discussion and Conclusion

The presented work is part of a project, in which a Digital Twin for energy efficiency and sustainability in the manufacturing sector is developed. The focus of this paper is the analytical level of the Digital Twin, which addresses the gap in research of Digital Twins with sustainability as optimization parameter. The work-in-progress presents results for a case study, using a simplified process with two machines, in order to provide a proof-of-concept for the planned contributions. These are the development of a framework for the assessment of sustainability of a real production process, with respect to the measured consumption of the different resources, and a concept for privacy-preserving data sharing within the supply chain.

The framework is presented and the results of the case study successfully identified significant energy efficiency and sustainability optimization possibilities, achieved with very simple modeling. It shows how for machine B, including only electrical energy in the energy consumption assessment neglects approximately 18% of the total consumption. Additionally, it demonstrates how the simple visualization allows to identify that machine A consumes 40% of the maximal operating power during the weekends, when nothing is being produced. This proves the potential of Digital Twins for the presented problem. The proposed framework is easy to deploy and generic, which means that it can be used in many different manufacturing processes. Visualization of the results of the model is user-friendly, which is crucial for the objective of supporting decision-making for production planning. Finally, ordinal encoding of categorical data represents a good possibility of privacy-preserving data sharing for the simple case presented.

Future work includes the installation of the remaining sensors for the relevant measurement points in the process, which include other relevant resources (water, gas, heat and manpower). Additionally, once enough historical data is gathered, more complex machine learning models will be trained and deployed, in order to have real-time process control, which is the ultimate goal of a Digital Twin. Similarly, once enough data is available, better privacy-preserving techniques will be used, such as data synthesis, in order to preserve the data utility for the models, without risks of data leakage.

References

[1] Lee, J. & Bagheri, B. Kao, H. (2014). A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. SME Manufacturing Letters.

[2] Dalibor, M. Jansen, N. Rumpe, B. Schmalzing, D. Wachtmeister, L. Wimmer, M. Wortmann, A. (2022). A Cross-Domain Systematic Mapping Study on Software Engineering for Digital Twins. Journal of Systems and Software.

[3] Kritzinger, W. Karner, M. Traar, G. Henjes, J. Sihn, W. (2018). Digital Twin in manufacturing: A categorical literature review and classification. IFAC-PapersOnLine.

[4] Biesinger, F. Meike, D. Kraß, B. Weyrich, M. (2019). A Digital Twin for production planning based on cyber-physical systems: A Case Study for a Cyber-Physical System-Based Creation of a Digital Twin.

[5] Liu, C. Vengayil, H. Lu, Y. Xu, X. (2019). A Cyber-Physical Machine Tools Platform using OPC UA and MTConnect. Journal of Manufacturing Systems.

[6] Lee, J. Azamfar, M. Singh, J. Siahpour, S. (2020). Integration of Digital Twin and Deep Learning in Cyber-Physical Systems: Towards Smart Manufacturing. IET Collaborative Intelligent Manufacturing.

[7] He, B. Bai, K. (2020). Digital Twin-based sustainable intelligent manufacturing: a review. Advances in Manufacturing.

[8] Tao, F. Sui, F. Liu, A. Qi, Q. Zhang, M. Song, B. Guo, Z. Lu, S. Nee, A. (2018). Digital Twin-driven product design framework. International Journal of Production Research.

[9] Qi, Q. Tao, F. Zuo, (2018) Digital Twin service towards smart manufacturing. Procedia CIRP 72:237–242

[10] Botkina, D. Hedlind, M Olsson B (2018) Digital Twin of a cutting tool. Procedia CIRP 72:215–218

[11] Henning, S. Hasselbring, W. (2021). The Titan Control Center for Industrial DevOps analytics research, Software Impacts

[12] Henning, S., Hasselbring, W., Burmester, H. et al. (2021). Goals and measures for analyzing power consumption data in manufacturing enterprises.

[13] GDPR: General Data Protection Regulation. Reform of EU data protection rules. European Commission. (2018)

[14] IBM, What is encryption? Data encryption defined. https://www.ibm.com/topics/encryption (accessed October 15, 2022)

[15] Imperva, What is data anonymization? https://www.imperva.com/learn/data-security/anonymization/ (accessed October 15, 2022)

[16] Cynthia Dwork and Nina Mishra and Kobbi Nissim, A Study of Perturbation Techniques for Data Privacy, https://theory.stanford.edu/~nmishra/cs369-2004.html (accessed October 15, 2022)

[17] Scikit-learn, class sklearn.preprocessing.OrdinalEncoder, https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.OrdinalEncoder.html

[18] Hands-on machine learning with Scikit-Learn, Keras and TensorFlow: concepts, tools, and techniques to build intelligent systems (2nd ed.). O'Reilly. (2019)