Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 104 (2021) 1292–1297

www.elsevier.com/locate/procedia

54th CIRP Conference on Manufacturing Systems

Concept and Architecture for Information Exchange between Digital Twins of the Product (CPS) and the Production System (CPPS)

Anna Vogt*[a,b], Ralph Klaus Müller*[b,c], Thomas Kampa[b], Rainer Stark[a,d], Daniel Großmann[c]

[a]*Technical University of Berlin, Chair Industrial Information Technology, Pascalstraße 8-9, 10587 Berlin, Germany*
[b]*Audi AG, Auto-Union-Str. 1, 85045 Ingolstadt, Germany*
[c]*Technische Hochschule Ingolstadt, Esplanade 10, 85049 Ingolstadt, Germany*
[d]*Fraunhofer Institute for Production Systems and Design Technology IPK, Pascalstraße 8-9, 10587 Berlin, Germany*

* Corresponding author. Tel.: +49-841-89-760734. E-mail address: anna.vogt@audi.de, Tel.: +49-841-9348-6521. E-mail address: ralph.mueller@thi.de

Abstract

The Digital Twin concept and CPS- and IIoT-based approaches are increasingly important topics concerning future Industry 4.0 architectures. They offer high potential for dynamical aspects in intelligent production planning and control as well as part traceability and documentation. Standardized information exchange is an upcoming requirement among the whole supply chain. This paper presents a concept for a Digital Twin architecture based on motor production in the automotive industry. The key aspect is an information exchange structure for Digital Twins of products and production systems that are combined using principles of Dynamic Aggregation.

Keywords: Digital Twin; Cyber Physical System; Production System; Information Model; Information Exchange; Intelligent Production Systems; Smart Production

1. Introduction

CPS, IIoT and Digital Twins are inevitable trends in smart manufacturing system research addressed in various concepts since they are considered to bear great potential for the realization of Industry 4.0 architectures. Yet, these concepts haven't succeeded in taking root in manufacturing practice on a large scale until now due to several reasons like inconsistent industrial and information technology throughout the supply chain but also because of a lack of detailed use cases allowing a profound proof and optimization of concepts. This work is based on an analysis of a valid use case scenario concerning part traceability and quality documentation in the automobile industry. Necessary information exchange in the context of this use case is designed as information flows between Digital Twins of the products being manufactured – in the following referred to as CPS – and the manufacturing system, also described as CPPS. This innovative concept taking into account principles of Dynamic Aggregation is implemented with a proof of concept architecture using real machine data from the mechanical production facilities of combustion motor parts. After a brief introduction in chapter 1, chapter 2 presents the state of the art followed by the concept development and implementation in chapter 3. Chapter 4 concludes this work with an outlook.

Nomenclature

CPS	Cyber-Physical System
CPPS	Cyber-Physical Production System
DT	Digital Twin
IIoT	Industrial Internet of Things
IPC	Inter Process Communication
MQTT	Message Queuing Telemetry Transport
OEM	Original Equipment Manufacturer
OPC	Open Platform Communications
PLC	Programmable Logic Controller

2212-8271 © 2021 The Authors. Published by Elsevier B.V.

10.1016/j.procir.2021.11.217

1.1. Motivation

Transparency of important product characteristics in general and the documentation of the manufacturing process in specific are major concerns in the automobile industry as well as in other industries with higher security aspects. Automobile OEMs are subject to increasing legal requirements concerning part traceability and quality documentation. A gap-free part traceability is currently required for emission related parts as demanded by the European Commission [11].
Moreover, quality management certification requirements regulate the documentation of quality checks for safety features [5] e.g., crack tests for brake discs. This product documentation is usually stored in individual plant databases and only accessible through an elaborate processes by the manufacturing companies in indicated situations. Especially when it comes to transparency throughout the manufacturing supply chain, in the current state it is hardly possible to access documentation on a component level in an efficient way.

1.2. Problem Statement

To ensure a cross-plant connection and dataflow, a reliable architecture is required for the building of a Digital Twin on component level. Storing of machining information is the basis for the concept. An intelligent mechanism must be developed to fulfill this task, because the structure of machining data of several suppliers and from different information sources in the production system might be different. Furthermore, the consolidation of this information in the sense of a Digital Twin is needed. Thus it appears that a secure data connection and a dynamic storing are requirements to solve the main problem of the information exchange between *Digital Twins of the Product (CPS)* and the *Production System (CPPS)*.

2. State of the Art

This chapter presents the state of the art concerning the main topics forming a basis for the developed concept: a definition of the Digital Twin (see section 2.1) as well as basics on Information Modeling Strategies and Dynamic Aggregation with respect to Digital Twins (see sections 2.2 and 2.3).

2.1. Digital Twin

Digital Twin technologies have been at the top of emerging technologies for the last few years [17], raising the need for a solid definition of what this term means in specific. According to the CIRP Encyclopedia a “Digital Twin” is defined as *“. . . a digital representation of an active unique product (real device, object, machine, service, or intangible asset) or unique product-service system (a system consisting of a product and a related service) that comprises its selected characteristics, properties, conditions, and behaviors by means of models, information, and data within a single or even across multiple life cycle phases.”* [21]
Eigner et al. state that the concept of Digital Twin does not necessarily imply that the whole amount of data of the digital representation of a (cyber-) physical product must be stored in one database [9], thus the crucial aspect is the ability of dedicated access to and connection of data stored in source systems via IIoT by means of Unique Identifier Numbers. Whereas classical business models for Digital Twins refer to the usage phase of products, the first potential in the product life cycle presents itself during the production phase when complex, intelligent products (CPS) and smart production systems (CPPS) get the chance to interact. Cyber-physical Systems are defined by their ability of intercommunication [10]. In order to detail the communication between Digital Twins during the production phase, a basic model is presented by Vogt et al. in 2020 describing a “Communication loop within twin sets CPS and CPPS” [24].
Different architectures and guidelines for digital twins have been standardized in the last years. In the standard of [13] of the Industrial Internet Consortium a digital twin architecture is examined. The authors focus on architectural evaluation criteria for the vision of future digital twins including e.g. Application Programming Interfaces (APIs) for data access or security aspects such as authentication modes.

2.2. Information Modeling Strategies

According to the authors of [19] and [20], information modelling is an important concept using standardized middle-ware communication technologies for the structured information exchange in complex production systems. Middleware with appropriate service protocols in automation has been established in the area of production automation since large amounts of data have to be handled and transmitted between different levels in production departments. Different guidelines like the VDI/VDE 2657 [23] describe the fundamentals of the current approaches. The main advantages of standardized middleware in automation are increased transparency, flexibility, maintainability, scalability of applications. The often used client-server and publish-subscribe approach brings improvements in interacting with different sources of information. Furthermore, standardized interfaces can be reused and the amount of different interfaces can be dramatically reduced in current and future production lines due to the simplification of development of distributed automation systems. An example for such a standard is OPC UA, which is used by the *RAMI 4.0 (Reference Architecture Model for Industrie 4.0)*. [23], [14]
Kutin et al. [15] present a concept for information links within designing processes of product and process production system. They highlight the necessity of the information interaction between design and production stages and processes.
The main concept presented by the authors of [16] is based on three approaches: the entity-relationship (ER) approach, the functional modeling approach, and the object-oriented (O-O) approach. The ER approach discusses the usage of entities and relationships to describe information requirements. The functional modeling approach is placed on specifying to the system functionality, whereas the O-O approach describes the identifying objects from the application domain first and then the re-

lated functions, methods and operations.
The authors of [20] deliver an approach for the *Dynamic Aggregation*. Due to the digitalization an increasing amount of information sources and sinks are integrated in a production environment. Aggregation addresses the challenge of managing different devices by integrating all information into a single point of entry. This concept is adapted in the next subsection with respect to Digital Twins.

2.3. Dynamic Aggregation with respect to Digital Twins

As previously described (see 2.1), Dynamic Aggregation is a suitable solution to address the upcoming challenges of the realization of Digital Twin architectures in IIoT-based smart production.
The authors of [18] investigate a connection architecture for a Digital Twin over six layers with focus on a demonstrator for a typical manufacturing component. A production cell-based scenario focusses on layered Digital Twins with levels of devices and sensors, data sources, local data repositories, IoT gateways, cloud-based information repositories as well as emulation and simulation. In this scenario, the aggregation concentrated on a Digital Twin that calculates an average power consumption of a combination of devices in a manufacturing process [18].
Another approach concerning a digital process twin is described by [6]. In this elaboration, they focus on a milling process and develop an information model that is the basis for certain calculations and analyses of the process. The aggregation lies in the acquisition of the heterogeneous data which is converted into a structured model.
In the study of the authors of [1] main effects of networks communication on manufacturing systems performance are outlined in the terms of process quality. Effects like delays on control signals, the packets loss and the measurement distortion are examined.
In relation to database approaches, the usage of NoSQL databases is getting more and more attention within industrial use cases. The authors of [7] create a cloud-based real-time monitoring architecture representing an integration of several machines through a direct attached system on the local side. The used database is InfluxDB as it is outlined that the core requirement of handling a large amount of raw data is sufficiently fulfilled.

3. Digital Twin Information Exchange Concept and Architecture

The following sections describe the path to a realized and tested industrial use case based on motor assembly in the automotive industry in an international production plant network. In order to realize a proof of concept implementation, the official company network system was utilized. Requirements for establishing the used cross-plant connection are not part of the realization of this concept.

3.1. Concept Development

In future industrial concepts, the connection between products, processes and resources will get more attention because legal and quality demands are increasing steadily. For that reason, an architecture that provides a first simple, possible solution is necessary to fulfill these requirements by tracking machine data and mapping it to a product during the manufacturing process.
The need for gathering manufacturing information is addressed by utilizing different kind of sensors directly in the machine. In the last years, this field has gained a lot more importance, as the intelligence is shifted partly into the sensor. Berger et al. [4] define in this context the term cyber-physical sensor system (CPSS) under the requirements of industrial applications of sensors. They classify the elements in a self-description model that can be used to get knowledge into the selection of a sensor for a specific use case. In our context, the information gathering starts at the PLC level, but it has potential to evolve in the future.
Most of today's products consist of many different parts that come from different suppliers along the supply chain. In the case of a motor assembly this could be the power electronics or the variable valve-control systems. The data during the machining process of a part can be acquired along the various production steps in the production system of a supplier through established standards such as *Open Platform Communications Data Access (OPC DA)* or in the future more widely established standards such as *Open Platform Communications Unified Architecture (OPC UA)*. OPC UA provides various modelling methods to structure information concerning semantics and syntax. A general concept architecture for data acquisition in the sense of a part-based Digital Twin is shown in figure 1.

In the shown architecture, the information source is the machining station of a supplier on the external side. In general, the data acquisition happens directly in the device fulfilled by function blocks connected to sensors. These function blocks are combined with objects in a communication protocol, for example OPC UA or MQTT. In the concept approach, the information model methodology or the structure of the MQTT communication object already lies in the considered entity. Current developments already cover the integration of customized information models directly in the device. For example, the CODESYS 3.5 SP17 [8] with the relevant software PLC covers the functionality of an own information model repository.
The data can be transferred over a local device which acts as the connecting part to a remote gateway in the transition layer. This layer has several security mechanisms such as a tokens and supplier-specific access roles. In modern, but mostly bigger companies this architecture already exists in certain ways for an interaction with suppliers. The adaption of such a structure for the concept is assumed, but it has to be validated for every specific use case. A data lake built upon a cluster of virtual machines is used as a basis for the internal side. According to the concept, every supplier has their individual virtual machine to perform the information transfer, where the design of the machine can be very lean according to the requirements and amount of part-based data. The information is forwarded into

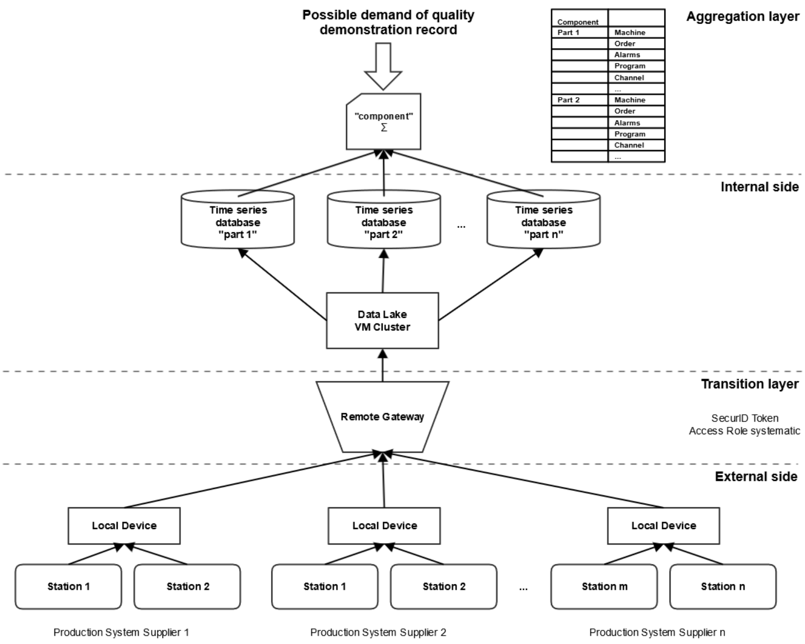

Fig. 1. General concept architecture for cross-plant Digital Twin information exchange

time series databases in a structured way. The structure can be outlined according to different classifications:

- **Unique part- or batch-based ID:** If possible, an identification number must be defined as early as feasible in the process for a part or batch. All machining information is stored in the database following this identification.
- **Workpiece carrier ID:** If possible, an identification number is used in combination with the workpiece carrier. The carrier and its ID follows the part through the production.
- **Production timestamp:** Specific machining information is provided with a timestamp. In combination with a production plan and the machine processes the classification follows.
- **Production entity:** The classification follows the stored information. An entity can be a production order, a machine identification or a machine manufacturing program.

After the classification inside the information storage databases, an aggregation layer follows. The single part information is aggregated to form a component information database, i.e. for a complete motor. A component dataset consists then of all machining information from different parts. Some possible examples are given in the following. *Machine* refers to location, connection and hardware information. *Order* describes a possible number coming from a production plan or program. *Alarms* include information about the unique failure ID, the type of alarm and a unique timestamp. The *Program* refers to the machining process, i.e. a NC-program whereas the *Channel* has information of the machine mode, i.e. hand-wheel-/ or teach-in-phases. The entities are combined with the concept of Dynamic Aggregation.

The requirements in the context for this method are the consistent structure of the acquired data and the implementation of mechanisms and rules to reach it for the considered suppliers. Therefore, a well-structured information model is needed. Otherwise dynamic aggregation gets more complex in terms of the configuration of an aggregation layer. Limitations in this context are the amount and detail of the data. Only the implemented combination between function blocks and information model objects in the PLC or tooling equipment can be used for the later digital twin representation. In general, a specific use case defines the structure of the information model, e.g. production data acquisition or energy consumption. The use case also defines the purpose of the acquired data, otherwise the data is collected in a data lake, but never used for analysis again. Furthermore, a continuous, confidential, low-loss and secure connection between the supplier and the enterprise cloud is needed for the concept. The concept doesn't observe the potential loss of data due to network connectivity and therefore outages and failures in the dynamic aggregation method.

In the next section, the architecture implementation is shown.

3.2. Concept Implementation

Based on the architecture concept developed in the previous step (see 3.1), this section presents the specific implementation as a living proof of concept which serves for validation purposes and is used to identify requirements and potential for improvement. This proof of concept utilizes the network infrastructure of the AUDI AG in order to transfer data from a machining center located in an international motor production to an industrial data cluster in Germany. Due to the usual durability of machines in a brown-field production line, the PLCs in operation require an IPC to receive data through OPC DA according to Fig. 2.

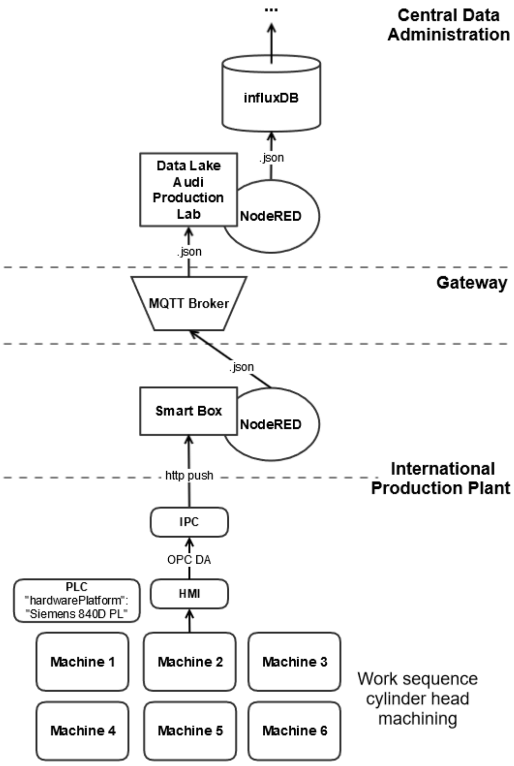

Fig. 2. Practical architecture on specific use case in the motor production

The IPC then sends the data to a smartbox via http push, where a Node-RED instance is running. The *Smart Box* is an edge device inside the production that fulfills several tasks regarding data transfer and acquisition, e.g., for bolt data documentation. In our scenario it realizes the connection to the upper gateway. The protocol MQTT is used in this work due to its predestination for lightweight data transmission [22]. Moreover, a firewall clearing already existed for MQTT data, which simplified realizing the proof of concept. Other protocols, i.e. OPC UA, can be used interchangeably with little customizations.
On the client-side, in the international production plant, we use Node-RED to publish the data with MQTT to the MQTT broker by HiveMQ. An Openstack cluster provided by the AUDI Production Lab, represented as *Data Lake* in Fig. 2, hosts the MQTT broker and all following components in separate virtual machines running Linux *Ubuntu 18.04 LTS* with custom specifications tailored to the requirements of the respective service.
The receiving end in Germany is another Node-RED MQTT client that subscribes to the dataflow and it parses the incoming data to store it in time-series databases. In our proof of concept, we use InfluxDB databases, which are a leading choice for storing time-series data due to their performance and wide range of functionalities [2] [3]. In order to improve speed and storage usage even further, no redundant data is stored in the databases. The parsing is performed in the following manner:

- Convert data from a JSON-string to JS-Object.
- *Optional*: Transform data to simulate a second machine.
- Discard unwanted data and split the incoming JSON into five flat JSON files.
- Assign the measurement accordingly to the expected Node-RED InfluxDB-node format.
- Convert the time to the expected nanosecond format required for InfluxDB.

This proof of concept is based on data of one machine only, another machine is simulated by means of making small customizations such as changing IP addresses and in Node-RED by modifying the JSON files. After successful parsing, we save the data of both machines in two separate InfluxDB databases, original and simulation in compliance with the modular architecture.
Another important aspect is the use of Python with the library *influxDB* to query data from the InfluxDB databases and for the subsequent Dynamic Aggregation. The Python-client receives all data in a given time span, resulting in a nested list in accordance to Dynamic Aggregation. The list can be exported into desired datatypes such as JSON-/ and .csv-files for further investigation and visual representation.
An exemplary excerpt from the output-file for a time span of 2 minutes is shown in Fig. 3.


```
[[
[{'name': 'alarms',
'columns': ['time', 'id', 'timestamp',
'type'],
'values': [['2020-12-03T00:38:58.987000Z',
'7022401',
'2020-12-03T00:41:33.873+01:00',
'Information']]}],
[{'name': 'channel',
'columns': ['time', 'handWheel', 'mode',
'teachIn'],
'values': [['2020-12-03T00:39:54.053000Z',
False, 'Automatic', False]]}]
]
[
[{'name': 'ncProgramm',
'columns': ['time', 'name', 'number'],
'values': [['2020-12-03T00:38:59.206000Z',
'AF0120_004_03_63.SPF',
17400.0]]}]
]]
```


Fig. 3. Exemplary excerpt from the output-file

4. Outlook

This research presents a Digital Twin information exchange concept and architecture that is tested for the use case of part-related production process documentation. Since it is successfully implemented as a proof of concept within a real industrial network, it is possible to draw first conclusions and discuss potential for improvement as listed:

- For an efficient implementation on a large scale, it is crucial that every product (or batch in early stages) is traceable by an ID at the very start of the production. This allows a gap-free documentation by means of Digital

Twin communication between products and production resources throughout the supply chain.

- The used database in the presented proof of concept is InfluxDB 1.83. As for now, InfluxDB 2.X is stable offering *buckets* within one database which makes separate databases for this use case obsolete.
- The MQTT data is parsed manually by creation of flat JSON-files. There is potential for automated parsing of the machine data. Another future-oriented possibility is to use the OPC UA standard and its *Extensible Markup Language (XML)* information model for automated parsing.

The proposed concept architecture is applicable to other use cases with little customization depending on the underlying business models, i.e. intelligent production planning and control as far as non-real-time applications are concerned. Also, an individual product carbon footprint from the production process seems feasible as machine energy data is well accessible.

The aspect of data security in Digital Twin architectures is a crucial aspect and therefore offers great potential for future work. In the work of Gehrmann and Gunnarsson [12] important security aspects concerning general digital twin architectures are outlined. The focus is on the analysis of a complete digital twin system scenario and proposal of an overall security architecture for such scenarios. In our context, as we use the enterprise network, the basic security topics concerning the usage of a MQTT broker were already examined by the security department. Nevertheless, for future deepening digital twin concepts in the considered real manufacturing environment, the proposed architecture can be analyzed and optimized according to security aspects such as robustness, consistency, confidentiality and integrity of the synchronization process.

Finally, even with the Industry 4.0 efforts of the last decade, the possibilities in a state of the art production are a point of start, but are still limited when it comes to Digital Twin applications. There is an obvious gap as far as the state of the research and especially the software vendors claims are concerned.

Acknowledgements

This project is funded by the AUDI AG. The presented approach was developed within the Chair for Industrial Information Technology at the Faculty of Mechanical Engineering and Transport Systems at the Technical University of Berlin and within the cluster "Digital Production" of the Research Institute AImotion Bavaria of the Technical University Ingolstadt.

References

[1] Papacharalampopoulos, Alexios, et al.; *The effect of communications on networked monitoring and control of manufacturing processes.* Procedia CIRP 41 (2016): 723-728.

[2] Bader, A., Kopp, O., Falkenthal, M.; *Survey and comparison of open source time series databases.* In: Datenbanksysteme für Business, Technologie und Web (BTW 2017)-Workshopband (2017).

[3] Balis, B., et al.; *Towards an operational database for real-time environmental monitoring and early warning systems.*

[4] Berger, Christoph, et al.; *Characterization of cyber-physical sensor systems.* Procedia CIRP 41 (2016): 638-643.

[5] Brückner, C., Bopp, R., Krauss, F.; *Qualitätsmanagement. Das Praxishandbuch für die Automobilindustrie.* Vol. 2.; 2019. p. 280, 281.

[6] Caesar, B., et al.; *Information Model of a Digital Process Twin for Machining Processes.* 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). Vol. 1. IEEE, 2020.

[7] Campos, J., et al.; *The Use of Relational and NoSQL Databases in Industrial Asset Management.* Engineering Assets and Public Infrastructures in the Age of Digitalization. Springer, Cham, 2020. 302-308.

[8] CODESYS GmbH; *Release CODESYS V3.5 SP17.* In: https://www.codesys.com/news-events/news/article/release-codesys-v35-sp17-2.html; 29.04.2021.

[9] Eigner, M., Detzner, A., Schmidt, P. H., Tharma, R.; *Definition des Digital Twin im Produktlebenszyklus.* In: ZWF 114; 2019. p 345-350.

[10] Eigner, M., Koch, W., Muggeo, C.; *Modellbasierter Entwicklungsprozess cybertronischer Systeme. Der PLM-unterstützte Referenzentwicklungsprozess für Produkte und Produktionssysteme.* Berlin, Heidelberg: Springer Berlin Heidelberg; 2017. p 15.

[11] European Commission.; *Commission Regulation (EU) 2017/1151 of 1 June 2017 supplementing Regulation (EC) No 715/2007 of the European Parliament and of the Council on type-approval of motor vehicles (. . .).* In: Official Journal of the European Union (C/2017/3521); 2017.

[12] Gehrmann, Christian, and Martin Gunnarsson; *A digital twin based industrial automation and control system security architecture.* IEEE Transactions on Industrial Informatics 16.1 (2019): 669-680.

[13] Harper, E., Christopher Ganz, and Somayeh Malakuti; *Digital twin architecture and standards.* IIC J. Innov.(November 2019) (2019).

[14] Heidel, R., Hankel, M., Döbrich, U., Hoffmeister, M.; *Basiswissen RAMI 4.0: Referenzarchitekturmodell und Industrie 4.0-Komponente Industrie 4.0.* Beuth Verlag, 2017.

[15] Kutin, Andrey, Vitalii Dolgov, and Mikhail Sedykh; *Information links between product life cycles and production system management in designing of digital manufacturing.* Procedia Cirp 41 (2016): 423-426.

[16] Lee, Y. Tina; *Information modeling: From design to implementation.* Proceedings of the second world manufacturing congress. Canada/Switzerland: International Computer Science Conventions, 1999.

[17] Panetta, K.; *5 Trends Emerge in the Gartner Hype Cycle for Emerging Technologies.* In: https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/; 2018.

[18] Redelinghuys, A. J. H., Kruger, K, Basson, A.; *A six-layer architecture for digital twins with aggregation.* International Workshop on Service Orientation in Holonic and Multi-Agent Manufacturing. Springer, Cham, 2019.

[19] Schmied, S., et al.; *Erstellung und Management von Informationsmodellen für bestehende Produktionssysteme.* at-Automatisierungstechnik 68.5 (2020): 325-336.

[20] Schmied, S., et al.; *Integration of Manufacturing Information via Dynamic Information Model Aggregation.* Vietnam Journal of Computer Science (2020): 1-18.

[21] Stark, R., Damerau, T.; *Digital Twin.* In: Sami Chatti und Tullio Tolio (Hg.): CIRP Encyclopedia of Production Engineering, Bd. 66. Berlin, Heidelberg: Springer Berlin Heidelberg; 2019. p. 1.

[22] Thota, P., Kim, Y.; *Implementation and comparison of M2M protocols for internet of things.* In: 2016 4th Intl Conf on Applied Computing and Information Technology/3rd Intl Conf on Computational Science/Intelligence and Applied Informatics/1st Intl Conf on Big Data, Cloud Computing, Data Science and Engineering (ACIT-CSII-BCD). IEEE, 2016.

[23] VDI/VDE 2657 Part 1:2013-01 (2013); *Middleware in industrial automation – Fundamentals.* VDI/VDE: www.beuth.de

[24] Vogt, A., Schmidt, P. H., Mayer, S., Stark, R.; *Production in the Loop – the Interoperability of Digital Twins of the Product and the Production System.* In: Procedia CIRP; 2020. p 3.