Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 100 (2021) 115–120

www.elsevier.com/locate/procedia

31st CIRP Design Conference 2021 (CIRP Design 2021)

Data administration shell for data-science-driven development

Andreas Löcklin*[a], Hannes Vietz[a], Dustin White[a], Tamás Ruppert[b], Nasser Jazdi[a], Michael Weyrich[a]

[a]*Institute of Industrial Automation and Software Engineering of University of Stuttgart, Pfaffenwaldring 47, Stuttgart D-70550, Germany*
[b]*MTA-PE Lendület Complex Systems Monitoring Research Group, University of Pannonia, Egyetem u.,10, POB 158, Veszprém H-8200, Hungary*

* Corresponding author. Tel.: +49-711-685-67305 ; fax: +49-711-685-67302. *E-mail address:* andreas.loecklin@ias.uni-stuttgart.de

Abstract

Data-science-driven development projects are increasingly gaining the attention of small and medium sized enterprises. Since SME are often lacking the necessary competencies in data science, cooperation with other companies or universities is required. The efficient handling of data is one of the main challenges in joint cross-enterprise development projects. Actual cost driver is the development of data by labeling and classifying the data by domain experts, which is very time-consuming and labor-intensive with large amounts of data. Furthermore, clearance processes also have a high potential to cause delays before data can be shared with project partners. Moreover, before the actual work can begin, it is often necessary to clean up and repair incomplete or noisy data. The concept of Data Administration Shell presented in this paper addresses the challenge of structured information sharing and information management in joint cross-enterprise engineering. The Data Administration Shell links data sets to information regarding data origin and already performed analyses including their results and program scripts. Adding relations and documentation facilitates the reuse of data sets for subsequent projects. For this purpose, the Data Administration Shell adapts the concepts serving the information sharing in the research field of manufacturing and Digital Twin. The evaluation of the Data Administration Shell was based on time-series measurement data from a production process optimization scenario. Here, the Data Administration Shell manages the data sets of time series data and facilitates the joint cross-enterprise engineering of data-driven solutions.

Keywords: Data Science; Development; Artifact Management

1. Introduction

Data-driven solutions can increase a company's success by enabling new or improving existing products and services. Utilizing data analysis and machine learning, for example, great success is achieved in the fields of computer vision, autonomous driving or smart manufacturing [1–5], and algorithms developed using data-driven methods enable the control of very complex processes. In turn, data-driven requirements engineering and data-driven decision making can be used to create a feedback loop to the user in order to refine products in a customer-oriented manner [6].

These successes in various application areas ensure that according to the empirical study among SME of Bauer et al. [7] more and more small and medium-sized enterprises (SME) are considering and preparing the use of data-science-driven development. Even though according to [7, 8] SME need longer to make use of new possibilities that data analytics offer, the survey among practitioners from Svensson et al. in 2019 [9] showed accordingly that data is valued differently in companies today, but the participants agreed that data will be a very significant asset in the future.

Successful data-driven projects require interdisciplinary teams that include business, technical as well as data science expertise. According to [7, 8, 10, 11] especially SME often cannot cover all fields of competence equally and data science projects require cooperation with other companies or universities.. This paper presents the Data Administration Shell (DAS) concept for information sharing and information management concerning data sets that are used in joint cross-

10.1016/j.procir.2021.05.019

enterprise data-driven development projects. The by the discussed DAS concept addressed problem of data set sharing with other companies or universities can help especially SME as they are more likely not to be able to cover all the competencies required for data-driven development projects with their own employees. The DAS concept builds on the concept of the Asset Administration Shell, which is well known in connection with Industry 4.0 and Digital Twins, and which rethinks information sharing in the area of manufacturing.

The remainder of this paper is structured as follows: Section 2 is dedicated to data types, data science process models, data management and a discussion of the Asset Administration Shell, which represents the theoretical basis for the Data Administration Shell (DAS). Section 3 then introduces the concept of the DAS. The evaluation of DAS is presented in Section 4 and finally a conclusion follows in Section 5.

2. Basics and Related Work

Before introducing and discussing the Data Administration Shell (DAS) concept, this section first examines what data sets in industry look like and what data types occur. Then, data science process models are referenced to explain the typical workflow of data-driven development projects. This is followed by a differentiation of DAS from other frameworks for data management. Finally, the Digital Twin paradigm and the Asset Administration Shell concept are discussed, which have great relevance in the field of information sharing in manufacturing and factory automation.

2.1. Data Types in Industry

On a basic level data differs, it is either numeric or categorical. Numeric data can be subdivided into continuous-valued, e.g. sensor data like temperature, and discrete-valued like the number of production cycles of a machine. Categorical data comprises of nominal data, e.g. information on a the origin of products that has no inherent ordering, and ordinal data that has inherent ordering, e.g. perceived skill level in a task rated from novice to advanced. [12–14]

These basic types of data are typically arranged in some form into either structured or unstructured data. The exact definition of structured data is up for debate. On the one hand you could argue that an image file is structured, because we know where each pixel is located and that in total it represents an image. [15] However, on the other hand, the actual content of the image is hidden to a normal computer program and cannot easily be extracted. Therefore, image and text data are unstructured. [16] We will follow the latter definition here. Hence, image data, which is often used for quality control in production environments, is unstructured. Structured time series data containing continuous sensor values are widely available in production and often used for tasks like predictive maintenance [17] and product quality optimization [18], because of the many sensors that monitor and enable modern manufacturing systems. When data is recorded it needs to be organized and contextualized, which can be done using nominal data (e.g. sensor id, manufacturing process), discrete values (e.g. production cycles, timestamp) and ordinal data (noise level). Data Science projects usually analyze static subsets of recorded data, while more data is still being recorded in the production facility. How these static and dynamic data subsets relate to each other is considered another form of data, that needs to be managed. Additionally, the data science process itself creates more data as well in the form of data cleaning scripts, experiments and results. These are mostly unstructured data in the form of programming code and visualizations scripts.

This section began by describing types of data, that exist in general and the different structural ways it can be organized in. It went on diagnosing, that data-science-driven projects in production environments typically contain all of the described types and different structures. This leads to the inherent complexity in the management of the data and the data science process itself [19].

2.2. Data Science Process Models

A process model helps with planning and the execution of projects. For data-driven development projects, the main steps are data gathering, data storing, data analysis and data evaluation according to [20–24]. These steps are defined in the two basic process models for data science projects that are the Knowledge Discovery in Databases (KDD, [20]) and the Cross-Industry Standard Process for Data Mining (CRISP-DM, [21]). All other process models in this field have a comparable structure and partially extend CRISP-DM that is in use for more than one decade. Extensions became necessary due to new developments and concepts like big and massive data, machine learning or collaborative working. Newly developed process models that further refine KDD and CRISP-DM are Team Data Science Process (TDSP, [22]), Analytics Solutions Unified Method for Data Mining (ASUM-DM, [23]) and Data Science Process Model (DASC-PM, [24]). The emergence of new process models shows the high relevance of improving data science project planning and manageability. In joint data-driven development projects, close cooperation reduces development efforts in the areas of data understanding and data preparation. An exchange of evaluation results in turn helps in the conception of new projects, for which application and business understanding is required.

2.3. Frameworks for Data Management

Data management aims at recording, storing, organizing and maintaining data throughout its lifecycle. The core of data management frameworks is a database, typically a relational database, which stores the data [25]. Functionalities that encompass the databases like defining its schema, transforming incoming data into the schema, performance monitoring and deleting data according to company rules are part of the data management as well. Through many, often complex, transformation functionalities, it is possible to transform all structured data of a company into so-called Data Warehouses like SAP Hana, Amazon Redshift and Oracle Exadata Machine [26]. Nevertheless, as discussed in Section 2.1, unstructured data is an important part of all available data. Typical Data Warehouse systems are not capable of handling these types of data. Hence, the concept of the data lake has been created, which is a single place where all data, structured and unstructured can be saved, ideally in its raw unprocessed form [25]. Data Lakes can only be hosted as a cloud system due to the extremely high memory and scalability demands. The information density of a data lake can be very low and extracting meaningful insights needs expert knowledge, typically a data scientist. Having a dedicated data science and cloud department for data management is a hurdle for medium-sized companies, whose core competence is not in this field and an impossibility for small companies [11]. It can thus be suggested, that there is a need for solutions that have flexibility comparable to a data lake, but usability and structure like a data warehouse, especially for small and mid-sized companies.

2.4. Digital Twin paradigm and Asset Administration Shell

A Digital Twin is understood as an container for all models and data of its physical counterpart that is always in sync with the physical world [27, 28]. As analyzed in [27–29], the Digital Twin paradigm receives a lot of attention especially in the application domain of manufacturing systems.

The Asset Administration Shell represents an formal, standards-based Digital Twin [29] developed by members of Platform Industry 4.0 [30]and, from 2021, then further developed by members of the new formed Industrial Digital Twin Association (IDTA, [31]). The Asset Administration Shell is a very flexible concept that is gaining practical relevance with the availability of reference implementations such as the AASX Package Explorer [32]. One of the current main use cases of the Asset Administration Shell is cross-enterprise information exchange [33]. The definition of common semantics for various types of applications is still work in progress and is achieved by defining submodels. These submodels represent metamodels that define what kind of information should be shared and how this information should be formatted. For the problem of joint cross-enterprise data driven development, the Asset Administration Shell concept is promising, but the focus is on managing physical assets rather than data sets. The concepts Asset Administration Shell and Digital Twin also include the storage and access to data of an individual physical asset. To access data, a physical asset is selected and there are links that point towards where the needed data is stored. As described in Section 3 and illustrated in Figure 1, the Data Administration Shell (DAS) adopts an orthogonal approach and is directly coupled with data and envelops it with additional information and links to development artifacts. It is envisaged that the DAS also stores links to physical assets that measured the data. As a result, the Data Administration Shell concept can be used without a Digital Twin of a particular asset and it can combine data of various sources. Therefore, the Data Administration Shell complements existing Digital Twins or Asset Administration Shells as illustrated in Figure 1.

3. Data Administration Shell concept

The Data Administration Shell concept represents a lightweight and with open source software realizable solution for sharing data sets in joint development projects. It does not exclusively address SMEs, but SMEs in particular usually have only a rudimentary or no infrastructure for managing data sets and at the same time face the major challenge of needing to share data with external experts. Sharing data sets with other companies or universities is often the only possibility for SME to perform data driven development projects as SME often are not able to cover all competencies needed for such projects with their own employees. [7, 8]

The core of the Data Administration Shell concept is the information model shown in Figure 3. This metamodel defines what additional information besides a data set should be shared and also kept updated over the entire life-cycle of a certain data set. In order to obtain a deployable solution based on the information model, the open source software AASX Package Explorer [32] shown in Figure 4 was used. It allows the creation and management of a DAS for a data set. The realized DAS is a single .aasx file that can be easily shared and saved along with the data set it manages. Collaboration in data driven development projects with a DAS for data management is illustrated in Figure 2, work on a data set is documented in its DAS. In the following, first the underlying information model and then the realization will be discussed.

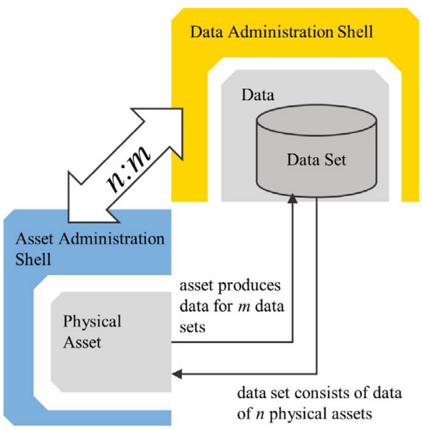

Figure 1: Asset and Data Administration Shell complement each other

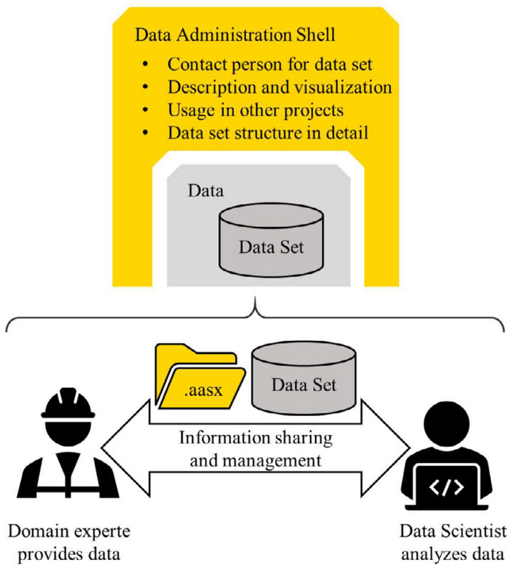

Figure 2: Sharing of a data set together with its Data Administration Shell

Data is precious and contains various causal relationships to other data. Therefore, data can be used in different projects and the Data Administration Shell (DAS) facilitates the reuse of already cleaned and labeled data sets. The DAS manages data over its individual lifecycle, from generation to project deployment to final deletion. The DAS serves as kind of a central guidepost to data and developmental artifacts. As shown in Figure 1, the DAS focuses on data and links to physical assets if necessary and is related but orthogonal to the concepts of Digital Twin and Asset Administration Shell. The machines that generated the data are also important, but retrieving data via the respective machine is cumbersome, especially with compound data sets. It is advantageous to manage data sets directly and, as part of the information stored for a data set, to link back to the machine that generated the data. The core of the DAS concept is the information model shown in Figure 3. It illustrates which meta-information should be stored for the management of data sets. For any data set, a contact person is linked and a general description of the data is given. If available, any kind of visualizations of the data set can also be linked to using the Data Administration Shell to allow for faster familiarization of project members with a data set. Furthermore, the DAS concept is not limited in this respect, but the authors recommend a project-related use. Data that is to be used for a specific project is managed in one or more DAS. For example, in Section 4, any projects using the data sets described in Section 4 would be logged by the DAS of the data set as part of a "Data Usage" description.

Clearance of data sets is another issue in joint cross-company data set sharing, which is not addressed by Digital Twin concepts or data management frameworks. For SME, data-driven development projects often require the cooperation with external experts as small and medium-sized companies have access to technical and application experts but lack competences in data analytics and data science [7, 8, 11].. Very early in cross-enterprise data driven projects the question "Which data can be given to project partners under which conditions" arises. To support here, the DAS stores not only development artifacts but also clearance information and thus facilitates the identification of suitable data sets that may be shared with project partners.

The essential part of the DAS concept is the definition of subsets to facilitate the work with data sets that are composed of many sub-data sets. As discussed in Section 2.4, Digital Twins are well suited for managing all data of a specific individual physical asset. Working with composed data sets is cumbersome as many Digital Twins and their respective interfaces come into play. For example, the "Smart Building" data set selected for evaluation purposes and discussed in Section 4 contains time series of temperature and power consumption curves, each of which is created as a separate subset with corresponding names and contact persons. For each subset, data properties and information on how the data was generated are also linked. Sharing a composed data set using a DAS is easier than sharing interface information of various Digital Twins, both from a technical as well as from a clearance and authorization perspective.

DAS can manage both sealed data sets and continuously growing data streams. Since raw data often needs to be cleaned up first or other transformations are necessary when dealing with data, each subset can refer to a parent and child subset. In the selected example, links between data subsets "OutdoorTemperature_1mHz_raw" and "OutdoorTemperature _1Hz_OutliersDeleted" can thus be mapped in the DAS.

As outlined in Section 2.3, there is a need for solutions that have flexibility comparable to a data lake, but usability and structure like a data warehouse. Large-scale commercial data warehouses can also achieve the objectives addressed by the DAS but lack flexibility. The DAS represents an easy to implement concept that facilitates the entry into the world of data-science-driven development. Especially for the first data-science-driven projects it is often not yet worthwhile to use commercial frameworks. With the experience gained in the first data science projects, a suitable framework can then be selected more specifically. The Data Administration Shell is therefore a bridging technology and a free, easy to use alternative. For implementation and evaluation, the publicly available tool "AASX Package Explorer" [32] is used as discussed in the next section.

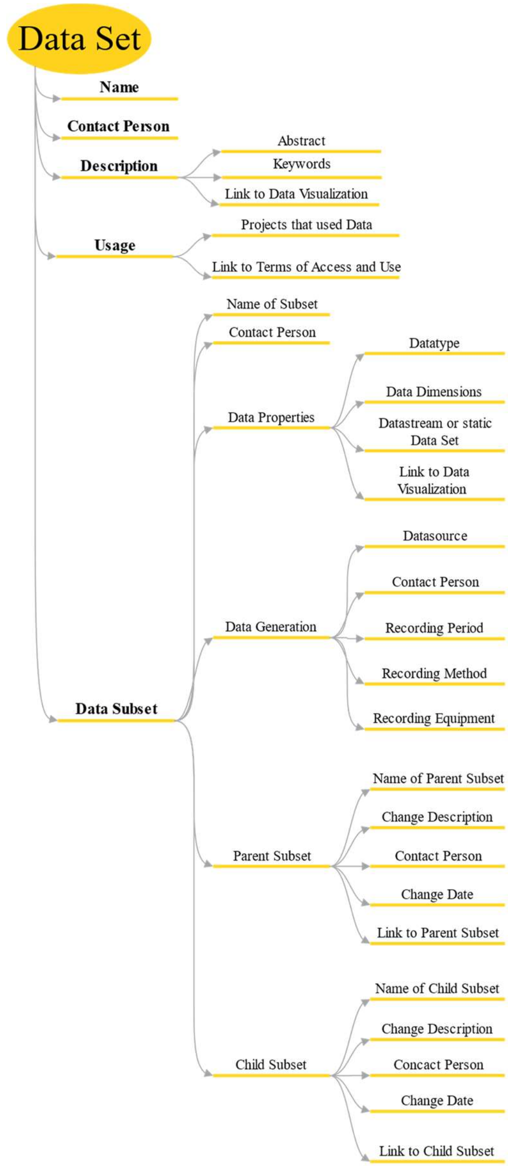

Figure 3: Information model of Data Administration Shell (DAS)

4. Evaluation of Data Administration Shell

To evaluate the Data Administration Shell (DAS) concept, the information model illustrated in Figure 3 was implemented utilizing the tool “AASX Package Explorer” [32], which is open source and licensed under Apache License 2.0 [34]. This tool is intended for the implementation of the in Section 2.4 described Asset Administration Shells. Basically, it is an XML file explorer that can be used for many purposes, including the realization of the Data Administration Shell.

For evaluation, a DAS was used to manage data sets as part of an analysis of the influence of temperature changes on the production process of a small company. The project partners lack expertise in the field of data science and therefore cooperate with universities. To keep track of which data can be shared, there is a lack of infrastructure to manage the data sets.

This is where the DAS comes in, acting as a guidepost between data sets and development artifacts. For each data set, a DAS is created in which contact persons and clearance levels are linked. Descriptive keywords are stored for each data set to ensure later browsability of the company's data sets. In addition, it is recorded which project partners have already worked with a data set and who has authorized the release of the data and for what purpose. Since the data sets themselves often consist of many values recorded in parallel from different sensors, the actual content of such composite data sets is described via several subsets. For each subset, the individual sensor that generated the data is documented. As tampered data, where some measured values are missing, is often not suitable for data analytics or machine learning applications, the project partners needed to clean the data sets first, for example by adding average values for each mission value. In order to make it comprehensible what the original data are and with which methods they were cleaned, subsets can be linked to each other via parent (original) - child (derived subset) relationships. The DAS of a data set thus corresponds to a kind of logbook of a data set in which it is entered who did what and why.

Figure 4 shows a screenshot of a realized DAS of the “Smart Building” data set. The data set is publicly available from [35] and resembles the data sets managed by the project partners as it also holds weather, temperature and energy consumption data. Figure 4 shows the graphical user interface of the tool “AASX Package Explorer” that is used to realize the DAS of the “Smart Building” data set. The AASX Package Explorer graphical user interface shows the tree structure of the DAS information model in the center and on the right-hand side, information about the selected entry is shown. It is possible to link to websites, databases or local files. Furthermore, descriptions can be stored in different languages.

The DAS can be seen as a digital logbook for data. It stores general information about a data set. If the data set holds various kinds of data, the concept of subsets is used. In Figure 4, the data set “Smart Building” holds data about the outdoor temperature and power used by the fridge, both temperature and power used are managed as subsets. “OutdoorTemperature_1mHz_raw” also needed to be cleaned and there is a parent-child relationship with “OutdoorTemperature_1Hz_OutliersDeleted”. The cleaning process consisted of replacing outliers with a value generated by averaging the neighboring values. When sharing the specific data set with other partners, they can also use the cleaned data as well as check whether the procedure used for data cleaning is suitable for their application purpose. If for example averaging is not suitable for a specific application, the parent-child relationship between raw and cleaned data clearly identifies the original data.

If a data set is reused in other projects, existing DAS should be reused if possible as well. This increases the benefits of a DAS infrastructure. When using a DAS, three application phases can be distinguished. Before a new data-science-driven project starts, existing Data Administration Shells are used to search for reusable data sets that could be used in the course of a project. Next, at the beginning of a project, new data sets are created or existing data sets are extended and accordingly existing DAS are used or new DAS are created to track and link any changes. During and at the end of a project, scripts for visualization and transformation are linked in DAS and application knowledge is added to facilitate reuse of data sets. Reusing data sets helps to reduce data labeling effort, which requires domain expertise and is very labor-intensive.

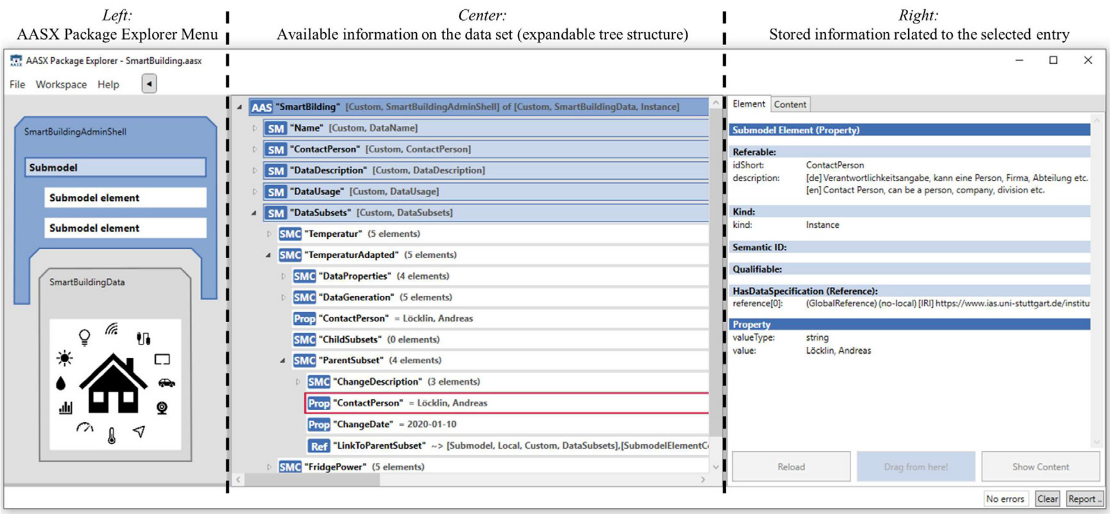

Figure 4: Screenshot of the Data Administration Shell of the “Smart Building” data set realized with the tool AASX Package Explorer

4. Conclusion

Small and medium-sized enterprises are increasingly interested in the data-driven improvement of their products and services. To carry out data-driven development projects, SMEs often rely on cooperation with external data science specialists from other companies or universities. In such joint cross-enterprise development projects, SMEs provide data sets to their partners. A major challenge here is the efficient management of data sets. In the process, the same questions arise repeatedly: Which data can be shared, which investigations have already been carried out on the basis of which data, how has the data been prepared - for this important information, that goes beyond data storage in the cloud or on hard drives, there is a lack of tool support. The Data Administration Shell concept presented in this paper serves as a kind of logbook for working with data sets and is a guidepost between data sets and development artifacts and can be easily realized using open source software. Thus, the Data Administration Shell supports efficient data sharing in cross-enterprise projects. The improved management of data sets increases their retrievability and reusability. By using the Data Administration Shell, costs can be saved because data does not have to be constantly re-recorded, released and prepared for use in data analysis or machine learning projects. Instead companies can more easily access well-documented data that has been used in previous projects.

Acknowledgements

This research has been funded by the Federal Ministry of Education and Research of Germany and the National Research, Development and Innovation Office of Hungary in the framework of EUREKA OMTRTLS, project number 01DS19039B. The responsibility for the content of this publication lies with the authors.

References

[1] Diez-Olivan A, Del Ser J, Galar D, Sierra B. Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0. Information Fusion 2019; 92–111.

[2] Sharp M, Ak R, Hedberg T. A Survey of the Advancing Use and Development of Machine Learning in Smart Manufacturing. Journal of manufacturing systems 2018.

[3] Wang J et al. Deep learning for smart manufacturing: Methods and applications. Journal of manufacturing systems 2018; 144–56.

[4] Wang D, Shang Y. A new active labeling method for deep learning. Beijing, China, 6 - 11 July 2014 2014 International Joint Conference on Neural Networks (IJCNN 2014), Beijing, China; 2014; 112–9.

[5] Maschler B, White D, Weyrich M. Anwendungsfälle und Methoden der künstlichen Intelligenz in der anwendungsorientierten Forschung im Kontext von Industrie 4.0. In Handbuch Industrie 4.0: Produktion, Automatisierung und Logistik. Wiesbaden: Springer Fachmedien Wiesbaden; 2020;

[6] Maalej W, Nayebi M, Johann T, Ruhe G. Toward Data-Driven Requirements Engineering. IEEE Softw. 2016; 1:48–54.

[7] Bauer M, van Dinther C, Kiefer D. Machine Learning in SME: An Empirical Study on Enablers and Success Factors. Americas Conference on Information Systems (AMCIS 2020) 2020.

[8] Coleman S et al. How Can SMEs Benefit from Big Data? Challenges and a Path Forward. Qual. Reliab. Engng. Int. 2016; 6:2151–64.

[9] Svensson R, Feldt R, Torkar R. The Unfulfilled Potential of Data-Driven Decision Making in Agile Software Development. In Agile processes in software engineering and extreme programming: 20th International Conference, XP 2019, Montréal, Canada, May 21-25, 2019, Proceedings. Cham, Switzerland: Springer Open; 2019; 69–85.

[10] Mittal S et al. A smart manufacturing adoption framework for SMEs. International Journal of Production Research 2020; 5:1555–73.

[11] Dorota Stadnicka, Andrea Bonci, Lorenzoni Emanuele, Grzegorz Dec, Pirani Massimiliano. Symbiotic cyber-physical Kanban 4.0: an Approach for SMEs.

[12] Rose D. Data Science - Create teams that ask the right questions and deliver real value. Berkeley, CA: Apress; 2016.

[13] Bruce P, Bruce A, Gedeck P. Practical Statistics for Data Scientists, 2nd Edition, 2nd ed.: O'Reilly Media, Inc; 2020.

[14] Döbler M, Großmann T. Data Visualization with Python. Birmingham: PACKT Publishing Limited; 2019.

[15] Aggarwal C. Data mining: *The textbook*. Cham, Heidelberg, New York, Dordrecht, London: Springer; 2015.

[16] Gharehchopogh F, Khalifelu Z. Analysis and evaluation of unstructured data: text mining versus natural language processing. Analysis and evaluation of unstructured data: text mining versus natural language processing, Baku, Azerbaijan; 2011 - 2011; 1–4.

[17] Benjamin Maschler, Hannes Vietz, Nasser Jazdi, Michael Weyrich. Continual Learning of Fault Prediction for Turbofan Engines using Deep Learning with Elastic Weight Consolidation.

[18] Lindemann B et al. Cloud-based Control Approach in Discrete Manufacturing Using a Self-Learning Architecture. IFAC-PapersOnLine 2018; 10:163–8.

[19] Polyzotis N, Roy S, Whang S, Zinkevich M. Data Management Challenges in Production Machine Learning. Proceedings of the 2017 ACM International Conference on Management of Data - SIGMOD '17, Chicago, Illinois, USA; 2017; 1723–6.

[20] Fayyad U. Advances in knowledge discovery and data mining. Menlo Park, Calif.: MIT Press; 1996.

[21] Shearer C. The CRISP-DM model: the new blueprint for data mining. Journal of data warehousing 2000; 5:13–22.

[22] Etaati L. Introduction to Machine Learning. In Machine Learning with Microsoft Technologies: Selecting the Right Architecture and Tools for Your Project. Berkeley, CA: Apress; Imprint: Apress; 2019; 3–14.

[23] Kristoffersen E et al. Exploring the Relationship Between Data Science and Circular Economy: An Enhanced CRISP-DM Process Model. In Digital Transformation for a Sustainable Society in the 21st Century: 18th. Springer; 2019; 177–89.

[24] Schulz M et al. DASC-PM v1.0 - Ein Vorgehensmodell für Data-Science-Projekte2020.

[25] Miloslavskaya N, Tolstoy A. Big Data, Fast Data and Data Lake Concepts. Procedia Computer Science 2016; 300–5.

[26] Hüsemann B, Lechtenbörger J, Vossen G. Conceptual data warehouse design: Universität Münster. Angewandte Mathematik und Informatik; 2000.

[27] Ashtari Talkhestani B et al. An architecture of an Intelligent Digital Twin in a Cyber-Physical Production System. at - Automatisierungstechnik 2019; 9:762–82.

[28] Sjarov M et al. The Digital Twin Concept in Industry – A Review and Systematization. 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vienna, Austria; 2020 - 2020; 1789–96.

[29] Löcklin A et al. Digital Twin for Verification and Validation of Industrial Automation Systems – a Survey. 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vienna, Austria; 2020 - 2020; 851–8.

[30] Plattform I4.0. [Online] Available: https://www.plattform-i40.de/PI40/Navigation/EN/Home/home.html. Accessed on: Feb. 28 2021.

[31] User organization "Industrial Digital Twin Association" founded. [Online] Available: https://www.vdma.org/en/v2viewer/-/v2article/render/52443301. Accessed on: Feb. 28 2021.

[32] admin-shell-io/aasx-package-explorer. [Online] Available: https://github.com/admin-shell-io/aasx-package-explorer. Accessed on: Feb. 28 2021.

[33] Tantik E, Anderl R. Potentials of the Asset Administration Shell of Industrie 4.0 for Service-Oriented Business Models. Procedia CIRP 2017; 363–8.

[34] Apache License, Version 2.0. [Online] Available: https://www.apache.org/licenses/LICENSE-2.0. Accessed on: Feb. 28 2021.

[35] T. Singh. Smart Home Dataset with weather Information. [Online] Available: https://www.kaggle.com/taranvee/smart-home-dataset-with-weather-information/. Accessed on: Feb. 28 2021.