Design of a Data Structure for the Order Processing as a Basis for Data Analytics Methods

Guenther Schuh, Matthias Blum
Institute for Industrial Management (FIR) at RWTH Aachen University, Aachen, Germany

Abstract–**Today, manufacturing companies are facing the influences of a dynamic environment and the continuously increasing planning complexity. Using advanced data analytics methods, processes can be improved by analyzing historical data, detecting patterns and deriving measures to counteract the issues. The basis of such approaches builds a virtual representation of a product – called the digital twin or digital shadow. Although, applied IT systems provide reliable feedback data of the processes on the shop-floor, they lack on a data structure which represents real-time data series of a product. This paper presents an approach for a data structure for the order processing which overcomes the described issue and provides a virtual representation of a product. Based on the data structure deviations between the production schedule and the real situation on the shop-floor can be identified in real time and measures to reschedule operations can be identified.**

I. INTRODUCTION

The future project Industrie 4.0 is the driving force to secure the competitiveness of high-wage countries such as Germany and to expand their leading position in production technology [1]. The fourth industrial revolution is mainly driven by the internet of the things and services. Using intelligent production systems, Industrie 4.0 enables the further development towards a resource and energy efficient as well as a highly productive and flexible production [2, 3, 4]. Thus, the collection and usage of data within the manufacturing environment is essential. In non-production industries, data analytics methods are already implemented to evaluate past incidences and consequentially predict future happenings. One example is given by the company *Yarra Trams* from Melbourne (Australia). It uses the data from 91.000 trams and other data sources in order to obtain insight and knowledge about disturbances, performance or passenger volume. The obtained data and the foresighted use of vehicles increased service quality to 99% [5]. The company *blue yonder* provides its retail customers a software as a service platform, by which forecasts based on large datasets are made. Hereby, client needs can be predicted more precisely [6]. As shown by these examples, the usage of data analytics methods has a high potential to increase the efficiency of value-adding processes.

The basis for analytics approaches builds a virtual representation of a product on the shop-floor - called the digital twin or digital shadow. The digital twin or digital shadow illustrates the virtual representation of data in manufacturing and represents the relevant data (e.g., order, geolocation and status). Similar to a flight data recorder the relevant data is stored in a time series format. Although, existing IT systems provide feedback data from the shop-floor, they lack a data structure which provides a virtual representation of a product in real-time [7, 8, 9]. The current state of planning systems can be summarized to:

- insufficient image of the current situation of the production in terms of feedback data
- unstructured data in IT-systems
- rigid structures and a lack of adaptability of planning systems
- no continuous check of the data in the planning system and the status of the order processes

Considering these problems predictions about the future state of the production and reliable statements about the current situation of an order are not possible. In order to overcome the described issue and successfully implement methods of data analytics inside the manufacturing environment this paper presents a data structure overcomes the described issue and provides the virtual representation of a product

II. MOTIVATION

In the area of data evaluation and analysis, business intelligence using models exist which can be summarized under the term analytics capabilities. In this paper, the term analytics in association with business intelligence is defined as follows: It is understood as a scientific process of mathematical-logical transformation of data to improve decision making [9]. Depending on the maturity level of analytical skills, four stages of data analytics can be differentiated: descriptive, diagnostic, predictive and prescriptive analytics (see figure 1) [10,11].

The first stage, called descriptive analytics, answers the question “What happened?”. Descriptive analytics aim at analyzing large amounts of data with the purpose of getting an insight and conclusion of what happened in the past. By analyzing interactions within the data diagnostic analytics pursue the question “Why did it happen?”. Predictive and prescriptive analytics support proactive optimization. The question “What will happen?” is covered by predictive analytics. Future behavior is predicted by methods of pattern recognition and the use of statistics. Prescriptive analytics form the last stage and answer the question “What should be done?”. Using optimization algorithms and simulation approaches concrete measures are suggested or even directly implemented [12].

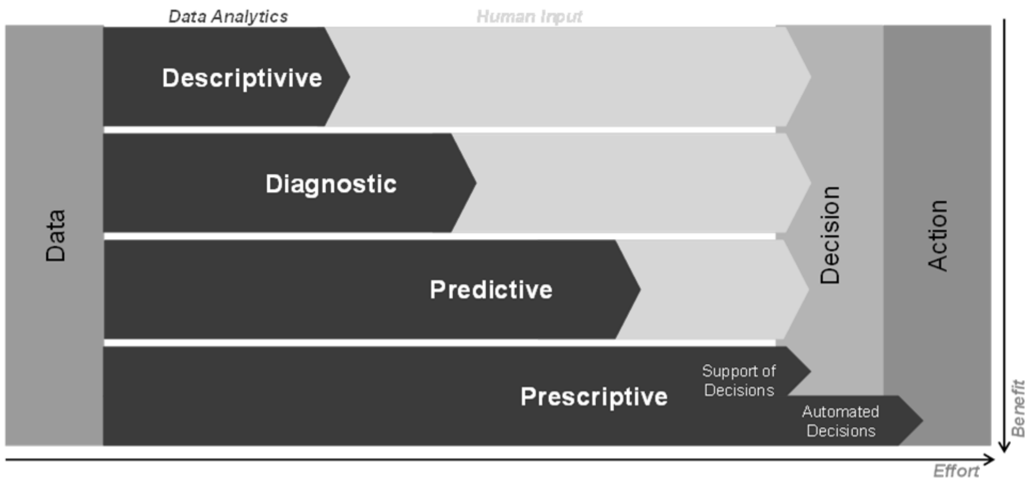


Figure 1: Four stage model of the decision making assisted by data analytics [11]

III. STATE OF THE ART

Although, several publications focus on approaches regarding data models which assist in planning and controlling the manufacturing process, a scientific investigation of a real-time representation of a product is only performed in very few research activities and not dealt with in detail. In the following section these approaches will be outlined.

Gröger designs an Advanced-Manufacturing-Analytics-Platform for a data-driven optimization of manufacturing processes. Based on an analysis of the manufacturing process he identifies the factors influencing the process and derives an analytic manufacturing process metamodel. Basis for this model are already existing process oriented metamodels and manufacturing modelling approaches such as the MES data model according to DIN EN 62264 (a standard which defines the enterprise-control system terminology), the value stream modelling or the virtual factory data model [14, 15, 16]. This approach enables to identity quality problems by a data-driven identification of root causes. As the focus is set on an optimization of quality problems, the order processing is not addressed [13].

Manufacturing execution systems (MES) are IT-systems which support the production management and are used for the process-oriented control and monitoring of manufacturing processes as well as resource and quality management. According to DIN EN 62264, the MES data base is formed by process model oriented, non order-related data of the manufacturing process. Those are material, personal and resource data which are required to form process segments. These segments refer to the manufacturing of specific products. [14, 17].

Terkaj conceives a virtual factory data model to integrate data from different domains (e.g., processes, products and resources). The purpose of the virtual factory data model is to enable the interoperability between the different software tools. It is a coherent, standard, extensible and common data model for the representation of factory objects. Providing real-time data and a focus on the order processing is not given [16, 18, 19].

Schuh designs a cyber-physical production control which aims at improving the short-term planning and long-term planning of the manufacturing resources. The basis for this approach builds a data structure which represents the relevant manufacturing data (e.g., shift plan, workstations and processes). Nevertheless, a discussion on real-time data of the order processing is not conducted [20, 21].

Another approach dealing with the modeling of the order processing is provided by Schmidt. The main focus of his work lies on the inter-company order processing and the development of an information model to improve the coordination between different companies in a supply chain [23].

Roos develops a data model for the intra-company order processing. Based on the Aachener PPS Model he defines the needed data for the tasks of the order processing. Although the data model is a central element of the work, providing a real-time representation of an order is not addressed [25].

IV. REQUIREMENTS

Based on the discussed approaches the following requirements have been derived which have to be met by a data structure for the order processing as a basis for data analytics:

- Data structure must illustrate relevant data of the order processing (R1)
- Appropriate for the single or small batch production (R2)
- Flexible Extension of the data structure to include other data (R3)
- Providing a real-time virtual representation of a product based on time series data (R4)

In Table 1 the reviewed approaches in regard to the derived requirements are compared.

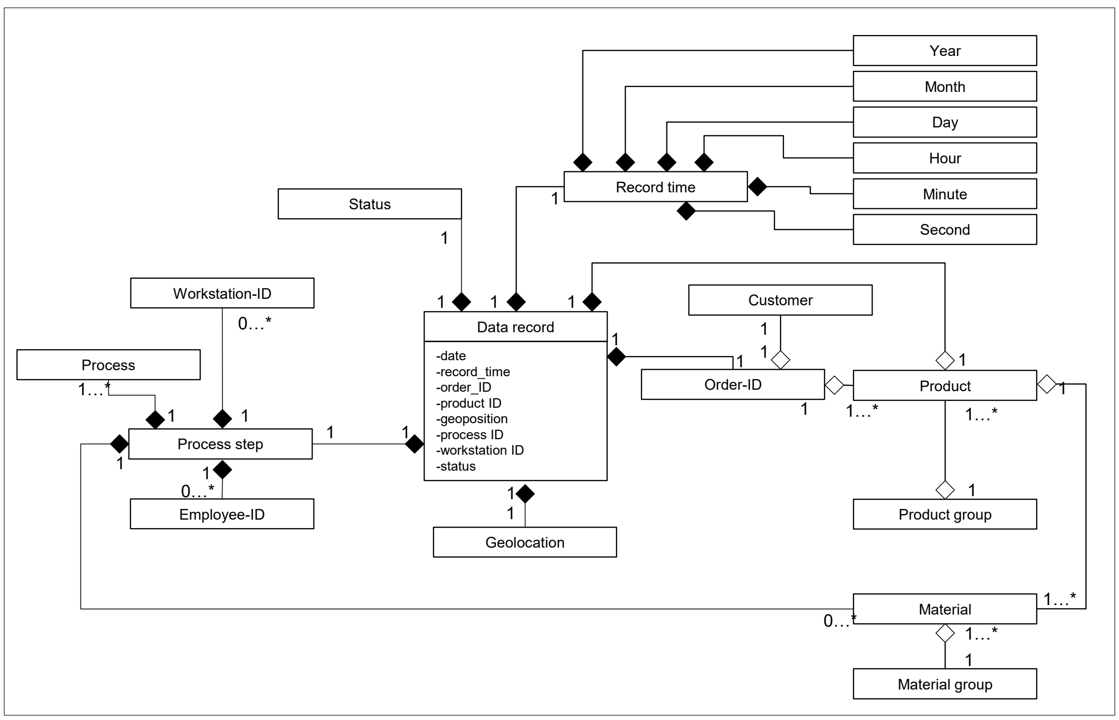
TABLE1: LITERATURE REVIEW REGARDING THE REQUIREMENTS

	R1	R2	R3	R4
Gröger	(yes)	yes	yes	no
MES-Datamodel	yes	yes	yes	no
Terkaj	no	(yes)	(yes)	no
Schuh	yes	yes	(yes)	no
Schmidt	(yes)	no	(yes)	no
Roos	yes	yes	no	no

The described approaches do not meet all requirements stated in the beginning. They provide a basis for a data model while lacking a full coverage of the whole order processing or a real-time virtual representation of a product. Based on the literature review and the derived requirements a data structure is presented in the next section.

V. DATA STRUCTURE AS A BASIS FOR DATA ANALYTICS

This section presents the derived data structure for the virtual representation of a product during the order processing as a basis for data analytics. For a comprehensive representation of an order on the shop-floor different aspects have to be considered. Besides tangible aspects (e.g. the product, the workstation, etc.) intangible aspects (e.g. process plans, the geolocation, the status, etc.) are needed. The data structure is primarily developed for single or small batch production of manufacturing companies. Thus, a special focus will be on linear and divergent production structures, the use of alternative resources and different operations and a semi-automated production with a high degree of manual process steps. Based on the production structure transport and temporary inventories will be considered. The derived data structure is presented in figure 2 and is described hereafter.



Date	Time	Order	Product	Geolocation	Workstation	Process	Status
10.01.2016	10:00:01	4711	A1234546788	90, 40	140	--	waiting
10.01.2016	10:00:02	4711	A1234546788	90, 40	140	--	waiting
10.01.2016	10:00:03	4711	A1234546788	90, 40	140	10	set-up
10.01.2016	10:00:04	4711	A1234546788	90, 40	140	10	set-up
10.01.2016	10:00:05	4711	A1234546788	90, 40	140	10	failure
…	…	…	…	…	…	…	…

Figure 2: Data structure for the order processing including example data sets

Time data specifies when the data was recorded and enables an entire data record from the start of monitoring until all processes are executed. Similar to an airplane's flight data recorder collected data is written inside a database in a specific time interval. As time data constitutes the leading characteristic among the collected data, corresponding data must refer to it. During order processing deviations from the schedule (e.g. geolocation, workstation and set-up time) may occur and can be detected in real-time.

Order data includes the attributes order and product. The item order assigns a unique ID to each order and enables the traceability of an order on the shop-floor. Furthermore, product ID is recorded to determine between different products of an order.

Based on the integration of new sensor technologies (e.g. real-time location system (RTLS) and radio frequency identification (RFID)) a live tracking of an order is possible. RTLS tags are applied to the product or the container and transmit the geolocation. Tracking the geolocation is necessary in order to ensure the routing of the order between two points. This enables to determine the current location of the order. Featured by the use of sensor technologies and a real-time routing, the status of an order between different steps in the working plan can be obtained.

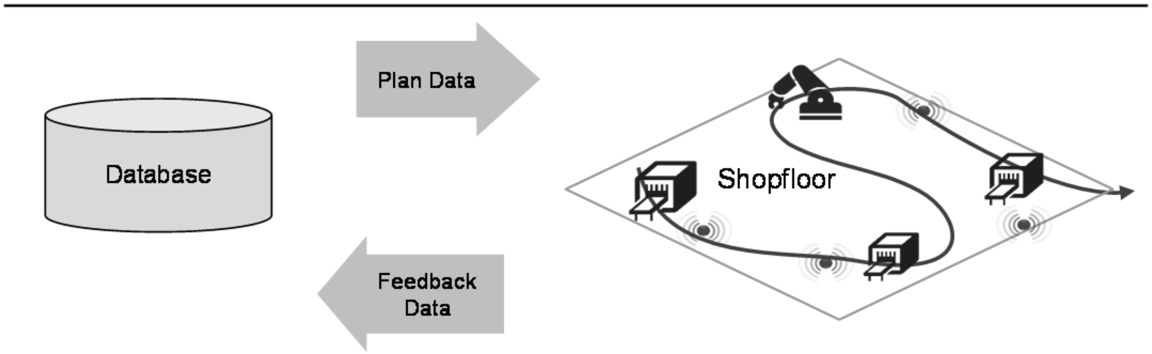
According to the understanding of a production process, the process consists of different process steps which are needed to produce a product. Items are tracked in relation to the working plan, e.g., the workstation and the process. This enables the user to determine on which workstations operations have been performed to complete the order. Based on the current operation the status of the order is logged.

To calculate time related data (e.g. absolute production time, set-up time, transition time) the status is needed. With this information conclusions about the current state of an order as well as ex post analysis are conducted. Based on the data record orders with the same production processes can be compared and reasons for deviations can be revealed.

VI. APPLICATION

Based on the results a prototypical implementation of the derived data structure to a demonstration factory is planned. Within the newly built Campus Cluster Smart Logistic the Demonstrationsfabrik Aachen (DFA) will provide a unique place to verify the results of the further research in a real production area of 1.600 square metres. The purpose of the DFA is to provide a place for research, training and real manufacturing. The application scenario includes the production of an electric go-kart. Using the data generated during production insights into current and past incidences are obtained (see figure 3).

Date	Time	Order	Product	Geoposition	Workstation	Process	Status
10.01.2016	10:00:01	4711	A1234546788	90, 40	140	--	waiting
10.01.2016	10:00:02	4711	A1234546788	90, 40	140	--	waiting
10.01.2016	10:00:03	4711	A1234546788	90, 40	140	10	set-up
10.01.2016	10:00:04	4711	A1234546788	90, 40	140	10	set-up
10.01.2016	10:00:05	4711	A1234546788	90, 40	140	10	failure
…	…	…	…	…	…	…	…



Date	Time	Order	Product	Geoposition	Workstation	Process	Status
10.01.2016	10:00:01	4711	A1234546788	90, 40	140	--	waiting
10.01.2016	10:00:02	4711	A1234546788	90, 40	140	--	waiting
10.01.2016	10:00:03	4711	A1234546788	90, 40	140	--	waiting
10.01.2016	10:00:04	4711	A1234546788	90, 40	140	--	waiting
10.01.2016	10:00:05	4711	A1234546788	90, 40	140	--	waiting
…	…	…	…	…	…	…	…

Figure 3: Using the data record to compare the plan vs. the real situation on the shop floor.

Collected data is compared to the plan data to allow observing changes in the order process over a time series. E.g., based on the real time data of an order, deviations in the processing can be uncovered in the moment when the processing time differs from the planned time. Therefore, plan data and feedback data are compared with upper and lower limit values.

VII. CONCLUSION AND FURTHER RESEARCH

In this research paper a data structure for the order processing is developed which aims at providing a virtual representation of a product during manufacturing, called the digital twin or digital shadow. The digital twin or digital shadow represents the relevant data (e.g., order, geolocation and status) in a time series format. After introducing the preconditions of the model the structural framework of a data structure for the order processing was derived. In order to determine the data structure, relevant system elements are derived describing the product and the processes. For enabling the implementation of the model, practical implications have been carried out. This data structure is the foundation for the use of data analytics methods in the manufacturing environment. Therewith, conclusions about past incidents and a real-time status of an order contribute to improve manufacturing processes. Further research is needed to substantiate the presented solution principles. Directions of further work include the development of measures to improve the data quality (i.e., plausibility and consistency checks and the use of redundant information provided by sensors), the handling of the geolocation and the development of a reference architecture.

ACKNOWLEDGMENT

The presented research is result of the Cluster of Excellence (CoE) on “Integrative Production Technology for High-Wage Countries” funded by Deutsche Forschungsgemeinschaft (DFG). Within the CoE “Integrative Production Technology for High-Wage Countries” several institutes at RWTH Aachen University are conducting research on fundamentals of a sustainable production strategy. The authors would like to thank the German Research Foundation DFG for the kind support within the Cluster of Excellence „Integrative Production Technology for High-Wage Countries.

REFERENCES

[1] Kagermann, H.; Wahlster, W.; Helbig, J.; “Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Securing the future of German manufacturing industry”, Berlin 2013.

[2] Koopmann, A.; „Industrie 4.0 ist gelebte Realität,“ in *VDI-Z Integrierte Produktion*, vol. 157, pp. 32–34, 2015.

[3] Sendler, U.; „Industrie 4.0. Beherrschung der industriellen Komplexität mit SysLM“. Xpert.press. Springer-Vieweg, Berlin 2013.

[4] Bauernhansl, T.; „Die Vierte Industrielle Revolution – Der Weg in ein wertschaffendes Produktionsparadigma,“ in *Industrie 4.0 in Produktion, Automatisierung und Logistik. Anwendung, Technologien und Migration*, T. Bauernhansl, Springer Vieweg, Wiesbaden 2014.

[5] IBM; “Descriptive, predictive, prescriptive: Transforming asset and facilities management with analytics”, Retrieved 04/14/2016 World Wide Web, https://static.ibmserviceengage.com/TIW14162USEN.PDF.

[6] BLUE YONDER; “Platform as a service”, Retrieved 04/14/2016 World Wide Web, http://www.blue-yonder.com/produkte/plattform.html.

[7] Schuh, G.; Potente, T.; Hauptvogel, A.; “Methodology for the evaluation of forecast reliability of production planning systems”, in *Proceedings of the 47th CIRP Conference on Manufacturing Systems (CMS) 2014,* ed.: H. Eimaraghy, Windsor, Canada: CIRP CMS April 2014.

[8] ElMaraghy H.; *Changeable and reconfigurable manufacturing systems*. London: Springer, 2009.

[9] Knabke, T.; Olbrich, S.: “Towards Agile BI: Applying In-Memory Technology to Data Warehouse Architectures”, in *Innovative Unternehmensanwendungen mit In-Memory Data Management. Beiträge der Tagung IMDM in Mainz*, ed.: W. Lehner., Bonn, Germany: IMDM 2011.

[10] Sherman, R.; *Business Intelligence Guidebook: From Data Integration to Analytics*. Waltham, MA: Morgan Kaufmann Publishers, 2014.

[11] FAIR ISAAC Coorporation; “Business Intelligence and Big Data Analytics: Speeding the Cycle from Insights to Action Four Steps to More Profitable Customer Engagement”. Retrieved 04/14/2016 World Wide Web, http://docplayer.net/3343658-Business-intelligence-and-big-data-analytics-speeding-the-cycle-from-insights-to-action-four-steps-to-more-profitable-customer-engagement.html.

[12] Stich, V.; Hering, N.; „Daten und Software als entscheidender Wettbewerbsfaktor“, *in industrie 4.0 magazin - Zeitschrift für integrierte Produktionssysteme,* vol. 1, pp. 8–13, 2015.

[13] Gröger, C.; *Advanced Manufacturing Analytics. Datengetriebene Optimierung von Fertigungsprozessen*. Lohmar, Germany: Josef Eul, 2015.

[14] Deutsches Institut für Normung (DIN); „DIN EN 62264 Integration von Unternehmensführung- und Leitsystemen“, Berlin: Beuth, 2014.

[15] Erlach, K.; *Wertstromdesign. Der Weg zur schlanken Fabrik*. Heidelberg, Germany: Springer, 2007.

[16] Terkaj, W., Pedrielli, G., Sacco, M.; “Virtual Factory Data Model”, in *Proceedings of the Workshop on Ontology and Semantic Web for Manufacturing,* eds.: D. Anastasiou, L. Ramos, S. Krima, Y. Chen, Graz, Austria: OSEMA, July 2012.

[17] Kletti, J.; *MES – Manufacturing Execution System. Moderne Informationstechnologie zur Prozessfähigkeit der Wertschöpfung*. Berlin, Germany: Springer, 2006.

[18] Terkaj, W.; Urgo, M.; “A Virtual Factory Data Model as a support tool for the simulation of manufacturing systems” in *Proceedings of the 3rd CIRP Global Web Conference on Production Engineering Research*, ed.: A. Caggiano, June 2014.

[19] Tolio, T.; Sacco, M.; Terkaj, W.; Urgo, M.; “Virtual Factory. An Integrated Framework for Manufacturing Systems Design and Analysis” in *Proceedings of the Forty-Sixth CIRP Conference on Manufacturing Systems (CIRP CMS 2013),* ed.: P.F. Cunha, Setubal, Portugal: CIRP CMS, May 2013.

[20] Schuh, G.; Reuter, C.; Hauptvogel, A.; Brambring, F.; Hempel, T.; „Einleitung“ in *Ergebnisbericht des BMBF-Verbundprojektes ProSense - Hochauflösende Produktionssteuerung auf Basis kybernetischer Unterstützungssysteme und intelligenter Sensorik,* G. Schuh, Apprimus-Verlag, Aachen, 2015.

[21] Schuh, G.; Hauptvogel, A.; Brambring, F.; Hempel, T.; „Entscheidungsunterstützung in einer cyberphysischen Produktionssteuerung. in *Ergebnisbericht des BMBF-Verbundprojektes ProSense - Hochauflösende Produktionssteuerung auf Basis kybernetischer Unterstützungssysteme und intelligenter Sensorik“,* G. Schuh, Apprimus-Verlag, Aachen, 2015.

[22] Schoth, A.; Quick, J,; Runge, S. „Gestaltung der PPS bei elektronischem Handel mit Produktionsleistungen“ in Produktionsplanung und -steuerung 2. Evolution der PPS, Günther Schuh and Volker Stich, Berlin, Heidelberg: Springer Vieweg, 2012.

[23] Schmidt, C.; „Konfiguration überbetrieblicher Koordinationsprozesse in der Auftragsabwicklung des Maschinen- und Anlagenbaus“. Aachen: Shaker, 2008.

[24] VDI-Fachbereich Technische Logistik; „VDI-Richtlinie: VDI 4400 Blatt 2 Logistikkennzahlen für die Produktion Blatt 3, Duesseldorf: Beuth, 2004.

[25] Roos, Erich; „Informationsmodellierung für PPS-Systeme. Ein Konzept zur aufgabenorientierten Systementwicklung“. Berlin, New York: Springer, 1992.