Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 72 (2018) 191–196

www.elsevier.com/locate/procedia

51st CIRP Conference on Manufacturing Systems

Substitution and Complementation of Production Management Functions with Data Analytics

Eduardo Colangelo [a,*], Torsten Kröger[b], Thomas Bauernhansl[a,c]

[a]Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Nobelst. 12, Stuttgart, 70569, Germany
[b]Institut für Anthropomatik und Robotik, Karlsruhe Institute of Technology, Engler-Bunte-Ring 8, Karlsruhe, 76131, Germany
[c]Institute of Industrial Manufacturing and Management IFF, University of Stuttgart, Nobenst. 12, Stuttgart, 70569, Germany

* Corresponding author. Tel.: +49 711/ 970 1912; fax: +49 711/ 970 1927. *E-mail address:* eduardo.colangelo@ipa.fraunhofer.de

Abstract

Complexity in production systems is steadily growing– one of the drivers is the customer's desire for personalization of products and services. Existing production management systems rely on deterministic functions. Given the diversity and varying influence of these functions, current methods reach their limits and may not meet future needs.
We introduce new opportunities provided by replacing and complementing these functionalities with data analytics. With a focus on production management and data-driven analysis we offer a flexible method that extends existing Industrie 4.0 technologies for several application areas (e.g., logistics, inventory management).

Peer-review under responsibility of the scientific committee of the 51st CIRP Conference on Manufacturing Systems.

Keywords: Data Analytics, Complexity, Production Planning and Control, PPC, MES, Industrie 4.0

1. Introduction

Production is entering the era of personalization. Aided by innovative technologies (like additive manufacturing) many possibilities are arising. Nevertheless, the biggest impulse comes from the customer side [1]. This means that manufacturing systems are, like never before, under a constant pressure to improve their performance and flexibility. In addition, there is a need for solutions that are ready to be implemented within the next few years.

In this context of growing complexity the following question arises: how can production systems be able to meet their logistic objectives while dealing with a completely unknown demand (in quantity and kind)?

In this paper, the specific problem of production management, as an area greatly affected by the increasing complexity (enforced by need for personalized products), is first explored. Then the proposed solution, the usage of data analytics as a way of dealing with this complexity, is introduced. Lastly, Industrie 4.0 is presented as the possible structure to implement it.

2. Complexity as a driver for data analytics

In recent years, scientists and entrepreneurs became aware that they need to abandon the idea of manufacturing systems being able to remain mostly unchanged, making only sporadic small adjustments. Instead, they need to adapt to steadily changing and not always clear environments. They started to consider the "complexity" of such systems.

In this context, complexity is defined as having four dimension: variety (the number of elements), heterogeneity (how different the elements are and the number of such differences), dynamic (how fast the conditions of the environment change), and opacity (the understanding of the elements, situations, and their relations; as well as their visibility) [2]. One example of a complex production system would be an enterprise with a large number of different products, providing a large supply chain (with many partners

2212-8271 © 2018 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 51st CIRP Conference on Manufacturing Systems.
10.1016/j.procir.2018.03.145

and low visibility of the specific requirements) and a fast development of new products.

It can be clearly seen, how the personalized production, with the possibility of the customer freely changing one or more features of products, has a great influence in increasing the complexity of production systems.

Different approaches exist to deal with complexity. Some focus on reducing and avoiding it. The focus of Fraunhofer IPA (and the one taken in this paper) is that the existence of complexity is not something that can be prevented. Two main reasons exists:

- It is not possible for an enterprise to fully control its complexity.
- A balance must exist between the internal complexity (i.e. variety of products, supply channels, etc.) and the external complexity (i.e. distribution of costumers, expected functionalities, etc.).

It then becomes clear that complexity is not something that can be eliminated. Doing so would mean facing the risk of not being able to adapt to the market development. On the other hand, the ability to correctly deal with complexity would then represent an opportunity [3]. Because of this, enterprises should try to "manage" complexity (i.e. create systems that are able to deal with the personalized demand in time and form) in order to profit from it.

Nevertheless, the complexity analysis is challenging. It essentially consists on finding the relations between factors (characteristics of complexity) using large amounts of data. In addition to every situation in each enterprise being different, the systems are continuously changing. Simple approaches, which rely on model-driven techniques, are unable to provide the required level of response [1].

2.1. Data analytics as an approach for complexity

In the area of information technologies, the necessity to analyze the ever growing amount of data led to the creation of the term Big Data. Many definitions exits, but the one chosen for the approach in this paper is that from Gartner: "Big Data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation" [4].

This definition handles Big Data as a mostly technical issue. But Big Data is at the same time a problem and an opportunity. This is the first similarity between complexity and Big Data. The second similarity lies on the challenges of Big Data and the ones of complexity analysis.

In order to deal with the problems (and opportunities) presented by Big Data, existing methods of data analytics were adapted to approach its three main characteristics: velocity, variety and volume (the 3 Vs). This means, that such methods need to deal with a great amount of fast changing data in different formats and from different sources.

The correspondence with the dimensions of complexity in production systems can, therefore, easily be found: velocity in Big Data corresponds to dynamic in complexity, variety in Big Data corresponds to heterogeneity in complexity, and volume in Big Data corresponds to variety in complexity. The concept of variety differs in Big Data and complexity.

The focus of data analytics, supported mainly by data mining and in part by machine learning, is to understand the underlying situation by processing the available data (data-driven approach). Or, in other words, to deal with the existing opacity (lack of visibility of the situation) by analyzing data. Then, the correspondence with the dimension of complexity of the same name is clear [5].

It is also important to refer to "available" data, because normally only a part of the data describing a situation can be found, accessed, and believed to be reliable. This is also true for production systems where both, internal (own production) and external (supply-chain) data, can be missing (i.e. lack of sensors, interfaces, etc.) or be false (i.e. production order confirmation being made out of time, rejected parts not being registered, etc.). These analyses must, therefore, be able to recognize patterns without having the whole picture, even when dealing with false input.

This problem is described in Big Data by means of an additional V: veracity. The correspondence with the opacity dimension of complexity becomes apparent, as false information affects the understanding of a situation and the corresponding decision making process negatively [6].

To the four considered Vs one more can be added: value. This V, understood as the degree to which an enterprise can profit from the data, also correlates to the expected benefit of managing complexity. In both cases, how to measure this potential gain is a challenge of its own [2].

It is then clear, how, due to the similarities between both challenges (Big Data and complexity), methods of data analytics can be applied to manage complexity in production systems.

2.2. The logistic target dimensions

Moreover, dealing with complexity in manufacturing systems also means improving the production logistics. How well this area works is measured by the logistic target dimensions: date compliance, lead time, performance, inventory, and costs [7].

These target dimensions compete constantly with each other [8]. For example, improving the date compliance may cause the costs to increase. That is why no simple mathematical optimization is possible [9]. The importance of each dimension depends on the particular situation and the interests of each organisation (i.e. prioritization of delivery dates).

Therefore, a coordination of the target dimensions is required. This is a difficult problem that occupies providers of PPCSs (production planning and control systems) and APSs (advanced planning systems) in their development of planning functionalities [10]. Such methods usually try to ponder the weight of each target dimension and, based on the expected progress of the production, deliver the best possible planning solution (local optimum).

With the additional increase in complexity, it can only be expected that the conventional methods and algorithms reach their limit. Data-driven analysis will, therefore, be necessary to deliver an adequate response, mainly be making use of its prediction capabilities, as explained in the next section.

3. Production management functions and the application of analytical methods

Much has been said and published in the last few years about data analytics and production. Some have proposed cloud-based approaches [11], a concept currently in development by automation giants like Bosch and Siemens. Nevertheless, these solutions fail in three aspects:

1. Further developments of IoT platforms mostly only cover solutions for specific use-cases (i.e. analyzing the performance of a machine). The problem isn't approached in an integrative way.

2. Logistic aspects of the production still aren't approached and, as a consequence of point 1, they can't be correctly approached.

3. Many solutions focus on creating the connection to the data sources, relegating the content of the problem itself to a secondary plane.

Data-based solutions for the area of production logistics must, therefore, consider the different existing functionalities and the way they interact. These are illustrated in Fig. 1.

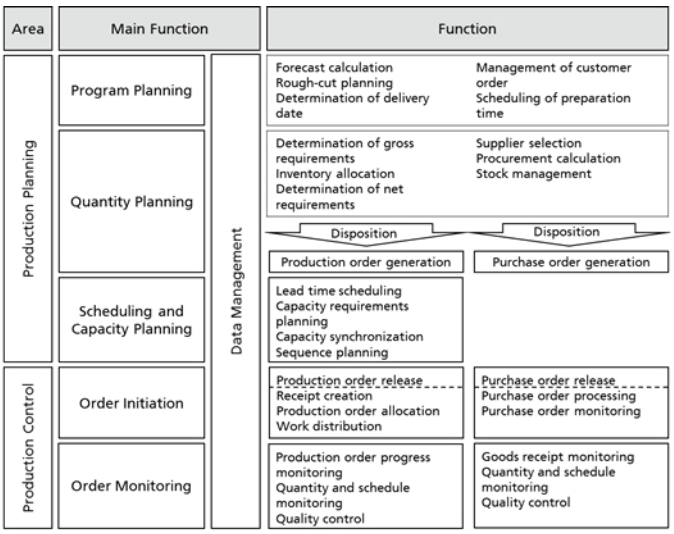

Fig. 1. Functionalities of PPCSs [12].

3.1. Current approaches and their limitations

A variety of methods have been developed to be used in the different areas of PPCS. Some can be simple, as the ones used in the production control to monitor that values are between determined limits. Some, on the other hand, are more complex. The area of production planning applies mainly formulas for the calculation of deterministic trends (some of them relying on stochastic models).

For example, working with the prediction of regular material requirements (times and quantities) leads to three types of formulas [13]:

- For constant requirements (i.e. regression).
- For requirements with some kind of trend (i.e. exponential smoothing).
- For requirements characterized by seasonality (i.e. Holt-Winters method).

The subsequent material requirements calculation (part of the disposition process) is a mere mathematical calculation aided by the usage of Gozinto graphs to model the relations between material requirements (primary and secondary requirements).

In problems of scheduling, a number of methods can be used for the batch size calculation. Some of them are DLSP (Discrete Lotsizing and Scheduling Problem), CSLP (Continuous Setup Lotsizing Problem) and GLSP (General Lotsizing and Scheduling Problem). They differ in the importance assigned to the setup time.

Similarly solved is the issue of calculation of purchase order quantities, with methods like UMSOQP (Uncapacitated Multi-Supplier Order Quantity Problem).

As stated in chapter 2.2, formulas used in the problem of order scheduling (capacity and sequence planning) usually use a weighted consideration of the constrains involved (which represent the logistic target dimensions). These can be further complemented with the application of simulation models (specially in APSs [14]).

Nevertheless, these approaches have clear limits.

When working in scenarios with big insecurities (for example, with totally unknown material requirements or with a fluctuating production flow), these formulas tend to try to increase the stochastic component. But this approach ends mainly in trying to calculate security buffers (in time and material). Such a take on the matter wouldn't be applicable in the personalized production, manufacturing unknown product variants while being unable to increase the production costs.

Furthermore, such formulas are limited in the number of elements they can consider. Either because the increase in their complexity would have a negative effect on their accuracy, or because too many variables would cause a counterproductive effect (a problem known as the curse of dimensionality) [15].

The usage of data analytics (based mainly on data mining and machine learning) to replace current formulas is then necessary to be able to work under production conditions of great complexity.

There are four types to be considered: Descriptive, Diagnostic, Predictive, Prescriptive [16].

Systems where the evaluation of the results is still mainly human will probably rely on descriptive and diagnostic analytics (used to understand a situation). They may also be used to assist in areas related to production, like quality and maintenance. Most of the work in last few years has been concentrated on this approach (i.e. production monitoring with business intelligence tools).

The usage of prescriptive and predictive analytics results more complex, as it requires an integrative vision (using domain knowledge) in order to create analytical solutions that take into account relevant and related factors (i.e. within order monitoring, the production times and the quality issues).

3.2. Towards prediction-driven services

As it can be seen in Fig. 1, while the control area is responsible for the order monitoring and small operative tasks; the main objective is the production planning, the creation of the production plan and its continuous adaptation. From the constitutive functions (program planning, quality Planning, scheduling and capacity planning) it can be derived that "planning" relies on predictions. They have to be performed based on the relations between many factors existing in a complex environment (unknown independent demands, availability of machines, availability of material, production times, etc.).

It is, for example, the case of the lead time calculation. This is composed by: processing time, setup time, waiting time, and transport time [7]. Each component is determined by many factors (most of them unrelated to each other). The prediction of the value is a problem which may even have to be divided in several sub-problems each using a different prediction model depending on the situation.

Planning systems in which this prediction models are fully integrated to the functionalities are the ultimate objective. For example: the sequence planning can be assisted by a prediction model of the production progress. This, in turn, can make use of the corresponding lead time prediction model for each personalized product.

Therefore, the relation between different functionalities (and their corresponding prediction models) also plays an important role.

This approach would also mean completely switching from a deterministic planning based on a known and foreseeable demand to a stochastic one, what allows working with unknown future requirements. Prediction models would allow to work not only with undefined working parameters (a big issue in personalized products), but also with the unknown effects in the whole working environments and manufacturing systems (i.e. the interaction between work stations). The coupling with prescriptive models complements the planning task by making the required decisions (i.e. when an order should be scheduled). Even the control tasks, normally focused on a shorter time horizon, can be covered with the usage of these models (i.e. monitoring the progress of a production order and triggering a planning task if necessary, what usually takes places within one work shift).

An improvement in the logistic target dimensions can be achieved, but a risk in the prediction exists. This can be reduced by using models with an appropriate accuracy (using the correct data in their creation) and by carefully choosing the areas and products where predictions are applied. Also, such model should be able to learn and improve itself.

The application of descriptive and diagnostic analytics in this area is also possible. On one side, as a way of complementing prediction models (i.e. by clustering of materials according to production parameters) and, on the other, as a support for remaining operative functions (i.e. diagnostic for analysis of quality issues and redesign of products or processes).

4. Challenges of data analytics

Though promising, the usage of data analytics is not without problems. Two main challenges are faced when applying data analytics to manufacturing systems:

- Data quality: Approximately between 70 and 80% of the effort employed to implement data analytics goes into the preparation of the data to be analyzed [17]. The data preparation also implies the usage of data analytics (i.e. classification for determination of missing values) [18].

- Additionally to data preparation, development, and implementation costs of data analytics are also considerable, even for relatively small projects [19,20]. Furthermore, data analytics projects are always a discovery process and, therefore, require going back and forth between the steps [21].

One key aspect is the usage of domain knowledge. As it can be seen in the Fig. 2, there are two approaches to create analytics: model-driven (top-down) and data-driven (bottom-up). The first one is mainly based on using domain knowledge (known relations between factors and existing models), but is less flexible. The second one tries to determine relations by looking at the data (classical machine learning approach), it allows the discovery of unknown behaviors, but the lack of domain knowledge makes it a long and costly process.

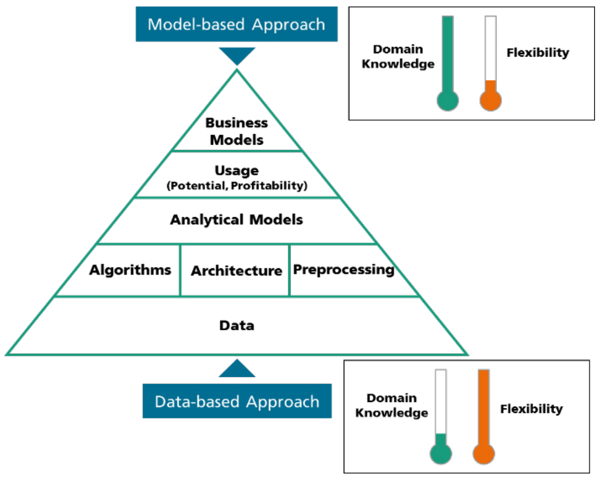

Fig. 2. Approaches for data analysis.

Of course, this division is not black and white, and both approaches profit from each other (i.e. by confirming the application of known models with machine learning). Domain knowledge helps reduce the effort and costs of analysis and data preparation in both approaches. Though it may hinder the flexibility, this depends on the way it is applied.

5. The structure of Industrie 4.0

The idea behind Industrie 4.0 is to create a new way of connection between the physical production environment through cyber-physical production systems, also known as CPPSs and the “intelligence” (the applications) that should control it.

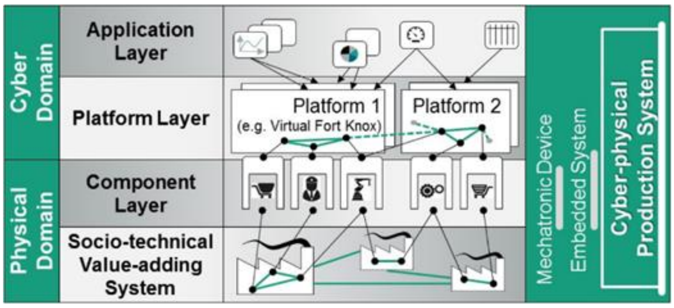

Fig. 3. The Industrie 4.0 structure [22].

Fig. 3 provides an overview of the main elements of Industrie 4.0. The upper level (the cyber domain) is based on the services approach. Here, every application is created, provided, used, and integrated as service. Services are both, functionalities that can work on their own and pieces of complex processes [23]. The platforms provide the basis through which these services can communicate with each other and with the physical domain.

In our case, the lower level (the physical domain) is where most of the data is to be found. The rest of the data comes from the processing results of each service.

The Industrie 4.0 structure can greatly assist the proposed approach [16]. The usage of data analytics as services (or microservices) provides:

- An easier way to integrate them with the functionalities of manufacturing systems. For example: a MES (manufacturing execution system) responsible for the sequence planning of the production can easily invoke the services responsible for the prediction of lead time. Moreover, because more functionalities of manufacturing systems are available as services, their integration with the data analytics services is both easier and more transparent. This helps reduce the implementation costs.

- A cost reduction through the way services are provided (also enabling new business models, like pay-per-use). The underlying Industrie 4.0 structure also helps reduce implementation costs (i.e. connection to data sources).

- An easier transfer of domain knowledge, as this will be embedded in the data analytics services and the way they are combined. This reduces the effort and costs involved in selecting an adequate analytical model (as a befitting one may already be available, requiring only minimal adjustments).

- An improvement of the data quality by eliminating the problem of schema integration (and thus reducing the preprocessing effort). As each component and application must have an appropriate interface for the connection to the platform layer, the formalization of the data models is forced.

- The possibility to flexibly integrate additional services for data cleansing (i.e. based on data analytics, ontologies, fuzzy-logic, etc.).

The Industrie 4.0 structure also provides the functionalities for data administration required by PPCSs. The connection to CPPSs can aid the order monitoring as a way to easily gather data [24].

6. Current applications

The Application Center Industrie 4.0 of Fraunhofer IPA provides an environment for developing and testing concepts and solutions of Industrie 4.0. One of the focuses is the creation of integrated systems that are able to deal with the requirements of the personalized production.

In this context, the proposed data analytics as services were introduced to complement and substitute several components of a production planning system for a prototypical manufacture of a personalized product. The utilized production systems consist mainly of machines for additive manufacturing (3D printers) and assembly stations.

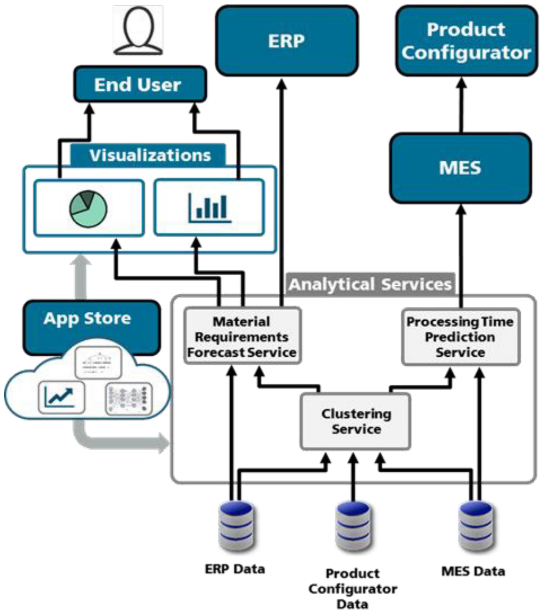

Fig. 4. Structure of demonstrator “Analytics Apps” and examples of analytical services.

A platform was created to allocate the developed analytical services, allowing the user to choose from a list of pre-configured analytical models (developed for production management applications) and implement them as services. It also enables the connection to both the data sources and the users of the performed analysis (ERP and MES). In contrast to this approach, available solutions on the market would require the realization of the connection work separately, not to mention the specific creation of the analytical models (as usually only the algorithms are distributed).

As it can be seen in Fig. 4, analytical services implemented within this demonstrator (called "Analytics Apps") allow, for example, the realization of material requirements forecasts (supporting the ERP), and the prediction of processing times based on past product configurations (supporting the MES). Both are done using models generated with the support vector machine method. These services are in turn complemented by a clustering service using the k-means algorithm. This creates groups of materials based on product and production features. Additionally to the ERP and MES, the used product configurator constitutes also a data source.

Other elements, like visualizations and a marketplace to distribute the analytical services, were also considered.

Benefits in the areas of application of the analytical services have already been observed (for example, very accurate prediction of processing times using few parameters, like machine type, volume and material). As it is an ongoing project, the effects are still being measured.

7. Conclusions and further work

As it can be seen, data analytics offer an adequate approach to overcome the problems of manufacturing systems that will arise with the increase in the level of complexity. Their application in areas which up until now haven't been sufficiently or correctly addressed (i.e. production logistic) show a lot of potential.

Industrie 4.0 provides both the required structure and the information backbone necessary for the proposed approach, especially as a way to integrate the developed solutions. As data analytics become more frequent in manufacturing environments, their integration in every aspect of production (to the user and the data sources) should be as easy and transparent as possible.

Risks exist, as the accuracy of the predictions performed is associated to the quality of the model used. The application of domain knowledge, the usage of abundant "good" data, and the constant improvement of the models (through self-learning), are ways to improve the accuracy and reduce risks.

Projects of Fraunhofer IPA, like the ongoing Application Center Industrie 4.0 and related ones, intend to explore and further develop this approach.

References

[1] Wehner D, Dangelmaier M, Hampel M, Paulus-Rohmer D, Hörcher G, Krieg S et al. Mass Personalization – Mit personalisierten Produkten zum Business-to-User (B2U); 2016.

[2] Colangelo E, Jäger, Jens, Paulus-Rohmer, Dominik, Schatz A, Bauernhansl T. Making the most of Big Data: Analytics-as-a-Service in the supply chain. In: Oduoza CF, editor. Flexible automation and intelligent manufacturing 2015 volume ii. Gloucester: The Choir Press; 2015, p. 522–529.

[3] Bauernhansl T, Schatz A, Jäger J. Komplexität bewirtschaften – Industrie 4.0 und die Folgen. ZWF 2014;109(5):347–50.

[4] Gartner. IT Glossary. [February 13, 2018]; Available from: https://www.gartner.com/it-glossary/big-data.

[5] Colangelo E, Pira A, Noack A, Bauernhansl T. Schritte zu Big Data in der Produktion. In: Zeitschrift für wirtschaftlichen Fabrikbetrieb, editor. ZWF (Zeitschrift für wirtschaftlichen Fabrikbetrieb); 2016, p. 851–854.

[6] Schuh G, Reuter C, Prote J, Brambring F, Ays J. Increasing data integrity for improving decision making in production planning and control. In: CIRP, editor. CIRP Annals: Manufacturing Technology; 2017, p. 425–428.

[7] Nyhuis P, Wiendahl H. Logistische Kennlinien. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012.

[8] Nyhuis P, Wiendahl H. Fundamentals of Production Logistics. Berlin, Heidelberg: Springer Berlin Heidelberg; 2009.

[9] Wiendahl H. Auftragsmanagement der industriellen Produktion: Grundlagen, Konfiguration, Einführung. 2011st ed. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012.

[10] Wiendahl H, Kluth A, Kipp R. Marktspiegel Business Software - MES - Fertigungssteuerung 2017/2018. 6th ed. Aachen: Trovarit; 2017.

[11] Gröger C, Niedermann F, Mitschang B. Data Mining-driven Manufacturing Process Optimization. In: Ao SI, editor. The 2012 International Conference of Manufacturing Engineering and Engineering Management, the 2012 International Conference of Mechanical Engineering. Hong Kong: IAENG; 2012, p. 1475–1481.

[12] Wiendahl H. Situative Konfiguration des Auftragsmanagements im turbulenten Umfeld. Zugl.: Stuttgart, Univ., Diss., 2002. Heimsheim: Jost-Jetter; 2002.

[13] Tempelmeier H. Material-Logistik: Modelle und Algorithmen für die Produktionsplanung und -steuerung und das Supply Chain Management. 5th ed. Berlin: Springer; 2003.

[14] Musselman K, O'Reilly J, Duket S. The role of simulation in advanced planning and scheduling. In: 2002 Winter Simulation Conference, p. 1825–1830.

[15] Breiman L. Statistical Modeling The Two Cultures. Statistical Science 2001;16(3):199–231.

[16] Colangelo E, Bauernhansl T. Usage of Analytical Services in Industry Today and Tomorrow. In: Westkämper E, Bauernhansl T, editors. Proceedings of the 49th CIRP Conference on Manufacturing Systems; 2016, p. 276–280.

[17] CrowdFlower. 2016 Data Science Report; 2016.

[18] Reuter C, Brambring F, Weirich J, Kleines A. Improving Data Consistency in Production Control by Adaptation of Data Mining Algorithms. In: Li Yea, editor. 9th International Conference on Digital Enterprise Technology (DET 2016): Intelligent Manufacturing in the Knowledge Economy Era, p. 545–550.

[19] Marbán O, Menasalvas E, Fernández-Baizán C. A cost model to estimate the effort of data mining projects (DMCoMo). In: Information Systems, editor. Information Systems; 2008, p. 133–150.

[20] pwc. Revolution Big Data; 2014.

[21] Fayyad U. From Data Mining to Knowledge Discovery in Databases. AI Magazine 1996;17(3):37–54.

[22] Landherr M, Schneider U, Bauernhansl T. The Application Center Industrie 4.0 - Industry-driven Manufacturing, Research and Development. In: Westkämper E, Bauernhansl T, editors. Proceedings of the 49th CIRP Conference on Manufacturing Systems; 2016, p. 26–31.

[23] Newman S. Building Microservices. 1st ed. Sebastopol: O'Reilly; 2015.

[24] Reuter C, Brambring F, Hempel T, Kopp P. Benefit Oriented Production Data Acquisition for the Production Planning and Control. In: Takata S, Umeda Y, Kondoh S, editors. Procedia CIRP: The 24th CIRP Conference; 2017, p. 487–492.