Linking data science to lean production: a model to support lean practices

Rossella Pozzi[a]*, Violetta Giada Cannas[a] and Maria Pia Ciano[a]

[a]*School of Industrial Engineering, LIUC University, Corso Matteotti 22, 21053 Castellanza, Italy*

*rpozzi@liuc.it

Linking data science to lean production: a model to support lean practices

Abstract. The literature discusses data science (DS) as a very promising set of techniques and tools to support lean production (LP) practices. DS could aid manufacturing companies in transforming massive real-time data into meaningful knowledge, increasing process transparency and product quality information and supporting improvement activities through data-driven decision-making. However, no attempt has been made in the literature to formalise the links between DS and LP practices. Thus, this study aims to overcome this gap by clarifying the DS techniques and tools that can support LP practices and how to apply them. This study employs a quantitative bibliometric method—specifically, a keyword co-occurrence network analysis—on a set of papers extracted from Scopus. The results obtained allowed the researchers to identify a set of DS techniques and tools that can support LP practices and to develop a model to guide their implementation based on the typical improvement implementation stages of the plan-do-check-act cycle. The model shows how to use DS techniques and tools in LP for: identifying areas for improvement and subsequent implementation (plan); enabling a better knowledge and process management (do); identifying/predicting potential problems and employing statistical process control (check); providing remedial actions and effectively applying process improvement (act).

Keywords: data science, lean production, big data analytics, keyword co-occurrence network, plan-do-check-act

1. Introduction

Lean production (LP) is as a pivotal strategy to enhance competitiveness (Bai et al. 2019; Belekoukias, Garza-Reyes and Kumar 2014; Bhamu and Sangwan 2014). It aims to reduce waste, improve efficiency (i.e. reduction in cost and lead time) and increase effectiveness (i.e.

quality enhancement of the manufacturing processes) (Womack, Jones, and Roos 1990; Jasti and Kodali 2015).

The literature demonstrated that the recent development of technologies and techniques for understanding phenomena via data analysis involved in data science (DS) can boost the outcome of traditional LP; also, DS outcomes can be empowered by LP practices, leading to better results than the ones that can be obtained from a purely technological adoption (Tortorella, Giglio, and Van Dun 2019). Thus, while implementing LP practices, companies can increase their outcomes by introducing new technologies, including DS techniques and tools. DS can improve LP practices potential (Rosin et al. 2020) by helping companies in achieving excellence through data-driven decision-making (Provost and Fawcett 2013; Wamba et al. 2015; Akter et al. 2016; Rejikumar et al. 2020). For example, real-time data acquisition and data analytics are very useful in dynamic value stream mapping (Wang et al. 2016; Meudt et al. 2017; Deuse et al. 2018; Phuong and Guidat 2018; Lugert, Batz and Winkler 2018). They improve employee commitment (Ciano et al. 2020) and process transparency, and support improvement (i.e. kaizen activities) through data-driven decision-making (Mayr et al. 2018).

However, even if DS potentially supports LP practices (Buer, Strandhagen, and Chan 2018; Rosin et al. 2020), only a few studies directly investigate this possibility (Antony, Gupta and Gijo 2018). Limitations in research emerge from the analysis of the articles addressing the study of the DS techniques and tools to support LP practices. The adoption of such techniques and tools has been investigated considering very few LP practices with respect to the entire set of practices studied by the literature and the implementation of DS techniques and tools in LP is still under-investigated. Narrow range of studies proposed or presented practical applications, and details on DS techniques and tools are lacking in most of the reported contributions, preventing further investigations.

Thus, the development of frameworks from the integration of technologies with LP is pointed as a valuable research direction (Kipper et al. 2020). Therefore, this paper aims to fill this gap by answering two research questions (RQs):

RQ1. What DS techniques and tools can support LP practices?

RQ2. How should DS techniques and tools be applied to support LP practices?

This first RQ objective is to define a set of DS techniques and tools that can support LP practices, by facilitating their implementation and boosting their results. Whereas the second RQ purpose is to define a model to guide the application of the DS techniques and tools that support LP practices. For achieving these goals, this study adopts a quantitative bibliometric method, specifically, a keyword co-occurrence network (Waltman, van Eck, and Noyons 2010), which it applies to the implementation of DS techniques and tools relevant to the key tenets of LP (i.e. production, process, inventory, workforce, supplier and customer). The results obtained allowed the researchers to identify a set of DS techniques and tools that can support LP practices, classified into a set of sequential activities to make DS applications successful. Second, the findings support the study in developing a conceptual model, grounded on the existing literature contributions, to guide the use of DS techniques and tools to support LP practices based on the typical lean implementation stages of the plan-do-check-act (PDCA) cycle.

The paper is organised as follows. Section 2 presents an overview of the background of this study. Section 3 includes a detailed explanation of the materials and the research method employed. Section 4 reports the results of the adopted bibliometric method. Section 5 presents the findings of this study and answers the two RQs, defining the links between DS techniques and tools and LP practices and providing an implementation model that encompasses and organises them. Contributions, limitations and directions for future studies are summarised in Section 6.

2. Background

2.1 Lean production practices

LP practices help companies to achieve higher efficiency and effectiveness and to strive towards operational excellence (Krafcik 1988; Womack, Jones, and Roos 1990; Belekoukias, Garza-Reyes and Kumar 2014; Bhamu and Sangwan 2014; Buer, Strandhagen, and Chan 2018; Tortorella and Fettermann 2018; Buer et al. 2020; Möldner et al. 2020). Their popularity has rapidly increased, moving from the restricted application in the automotive industry, where LP practices were first developed, to broader industries including transportation, process industries, and service sectors (Bhamu and Sangwan 2014; Jasti and Kodali 2015).

In the literature, LP practices have been often associated to LP 'bundles' (Shah and Ward 2003; Bevilacqua, Ciarapica, and De Sanctis 2017; Sancha et al. 2020). The bundles are 'made up of interrelated, internally consistent and even overlapping practices categories' (MacDuffie 1995, 6). The notion of bundles was introduced by Osterman (1994) and MacDuffie (1995) for human resource practices and then extended to LP practices by many scholars.

From the idea of the "lean enterprise", in which the company selects the best LP practices for functional areas and for external relations management (Karlsson and Åhlström 1996), Panizzolo (1998) developed a famous and recognised classification of LP practices into internal and external bundles that he called 'areas of intervention'. The recent paper by Bai et al. (2019) derives Panizzolo's framework and updates it by integrating and aggregating the results of subsequent papers that have dealt with classification, thus resulting in the classification of LP practices into 5 bundles: three bundles of internal LP practices (i.e. production planning and control, process technology and workforce) and two bundles of external LP practices (i.e. supplier and customer). Bai et al. (2019) considered all the previous literature on LP practices and represents the most recent and complete classification of LP practices in the literature. For these reasons, the classification of LP practices employed in this research is based on Bai et al. (2019) classification. Table 1 shows the classification and provides detailed description of each LP practice supported by literature.

Please insert here Table 1

Regarding the implementation of LP practices, no standard framework has been defined in literature (Bhamu and Sangwan 2014). About this topic, Marodin and Saurin (2013) report Value Stream Mapping, i.e., selecting a product family, mapping the current state, designing a future state and devising the implementation plan, as a common element among several studies on LP implementation and the use of the PDCA cycle for solving identified problems or addressing improvement opportunities in real cases. Accordingly, Watson and DeYong (2010) define PDCA as ‘the Japanese standard model for improvement and problem solving’, while Jones, Parast and Adams (2010) describe PDCA as ‘a well-established framework for process improvement’ and Linderman et al. (2003) suggested that ‘in case of process improvement the method is patterned after the plan, do, check, act (PDCA) cycle’. In the PDCA cycle, the ‘Plan’ stage considers the modelling and the objectives of the process. The ‘Do’ stage regards the management of the process. In the ‘Check’ stage, the organisation monitors and evaluates the process compliance with the Plan stage. Last, the ‘Act’ stage bases the related activities on the results of the check stage, improving the process (Dennis and Shook 2007; Liker and Morgan 2011; Marodin and Saurin 2013; Chiarini 2011).

2.2 Data science techniques and tools

In the recent years, the concept of ‘Data Science’ has represented an emerging trend in the field of business and industry. DS is defined as an interdisciplinary science that uses a mix of competences related to statistics, computer science, mathematics, data analytics, business and management, to quickly and effectively extract information and transform it into knowledge, by processing and interpreting massive volumes of data, and make better and more informed decisions (Vicario and Coleman 2020; Kamble and Gunasekaran 2020; Cao 2017; Provost and Fawcett 2013).

Nowadays, in industrial production, data-driven decision-making represents a key process to face the global competition and enhance productivity and profitability (Waller and Fawcett 2013; Hahn 2020; Raman et al. 2018), especially due to the emerging industry 4.0 technologies, which provide new ways to collect, store and analyse data (Kumar et al. 2018). These technologies are used not only in the isolated plants but also to share data in the supply chains (Hazen et al. 2014; Tan et al. 2015; Hazen et al. 2018). The rapidly increasing data are generated and shared continuously within and among companies (Agarwal and Dhar 2014; Sivarajah et al. 2017; Zhong et al. 2017).

Industrial big data are collected within the industry and are characterised by the ‘5V’ factors, ‘Volume’, ‘Variety’, ‘Velocity’, ‘Value’ and ‘Veracity’ (Addo-Tenkorang and Helo 2016; Baryannis et al. 2019). They are more available and easier to access and store than in the past (Waller and Fawcett 2013). To make the decision-making more meaningful, manufacturing companies should effectively analyse this massive amount of data. Hence, the role of DS becomes simultaneously fundamental and challenging (Xu, Xu, and Li 2018).

DS works down from the use of big data to describe, explain and predict phenomena (Jagadish 2015), and the data scientist is considered ‘a high-ranking professional with the training and curiosity to make discoveries in the world of big data’ (Davenport and Patil

2012). Hence, in manufacturing, DS is considered essential to better understand, manage and extract the hidden knowledge from big data, that is to perform big data analytics (Gupta, Modgil, and Gunasekaran 2020; Concolato and Chen 2017; Baesens 2014).

In the manufacturing literature, 'Data Science' and 'Big Data' are considered interchangeable terms (Jagadish 2015). However, DS is not limited to big data; it can be also applied to small data sets that contain valuable information (Van der Aalst and Damiani 2015). Additionally, according to the literature, DS requires companies to employ a structured procedure of implementation to extract valuable information from raw data. This procedure is characterised by a set of sequential activities required to implement DS successfully. Data analytics, or big data analytics, if applied to big data, is only one activity in this procedure.

Given the strong intertwining of DS with important concepts such as big data analytics, confusion exists in the literature about what exactly DS is and its role in serving business effectively (Provost and Fawcett 2013). Therefore, building on the previous literature contributions in this field (Donoho 2017; Vicario and Coleman 2020; Farooqui et al. 2020; Vitari and Raguseo 2019; Raman et al. 2018; Van der Aalst and Damiani 2015; Davenport and Patil 2012; Concolato and Chen 2017; Bailer and Fisher 2020; Talia, Trunfio, and Marozzo 2015), we present in Table 2 examples of DS tools and techniques, which have been classified based on the activities that characterise DS implementation, in order to help the reader understanding of such a new and complex topic and to clarify the meaning that the present work gives to the DS activities, tools and techniques. This set of activities, tools, and techniques, built upon literature contribution, serves as a base for the following steps of the study.

Please insert here Table 2

2.3 Linking data science and lean production

Literature provides empirical evidence on the positive effects that the interaction of technology adoption, including DS techniques and tools, and LP practices can provide to companies, if it is used to create value for people and processes (Tortorella, Giglio, and Van Dun 2019). In fact, on the one hand, the adoption of lean methods and the application of a lean thinking is demonstrated to pave the way to the introduction of new technologies; on the other hand, new technologies can support LP practices and boost their outcomes (Agostini and Filippini 2019; Bittencourt, Alves, and Leão 2021). For this reason, this study is interested in understanding the link between DS and LP practices, especially analysing the role of DS in supporting LP.

The predominant category that specifically analyses the link between DS and lean is the one related to lean Six Sigma (LSS) (e.g. Bazrkar and Iranzadeh 2017; Antony, Gupta and Gijo 2018; Dogan and Gurcan 2018; Noori and Latifi 2018; Antony and Sony 2019; Belhadi et al. 2020; Park et al. 2020; Gupta, Modgil, and Gunasekaran 2020). However, LSS is not listed within the bundles of LP practices. It is a separate category that links LP and Six Sigma. The two strategies share the goal of achieving quality by reducing variation, coupled with a continuous improvement approach, and this often results in them being associated and applied as a unified strategy under the name "lean Six Sigma" (Gamal Aboelmaged 2010; Ciano et al. 2019). Nevertheless, Six Sigma differs from LP: in fact, it tackles undesired variations based on an own structured problem-solving method known as 'define measure analyse improve control' (DMAIC) and addresses quality management through advanced statistical methods (Shah, Chandrasekaran, and Linderman 2008; Tjahjono et al. 2010; Watson and DeYong 2010).

A narrow range of studies proposed or presented practical applications of DS to specific LP practices. Ito et al. (2020) adopt simulation as a data analytics tool applied to field data in a real-time Andon system, supporting ‘visual management of production control’. Abd Rahman et al. (2020) apply simulation as a data analytics tool to generate overall equipment effectiveness (OEE) values as the ‘feedback on performance metrics practice’ of a production line. Shahin et al. (2020) propose a cloud Kanban system, supporting ‘pull production’. The cloud computing combines production data gathered from the field and results from an ant colony-based simulation serving as data generator, based on process rules and displays jobs and activity progresses. ‘Pull production’ is also addressed by Deuse et al. (2018), who propose the improvement of traditional Kanban using machine learning methods. Clustering is used in the identification of product families with similar routing and work content to deal with value adding variables, discovering multivariate connections and patterns in manufacturing data. Stojanovic and Milenovic (2018) present an approach that exploits process mining, namely deriving a process model from process data, and analytics tools to check variations from the obtained model and real time data and apply machine learning techniques to allow ‘root cause analysis for problem solving’. Kutschenreiter-Praszkiewicz (2018) propose using the machine learning techniques of neural networks and decision trees for connecting elementary motions into activities and setting manual task time standard to support the streamlining step of the ‘setup time reduction’. Sarkar et al. (2020) propose the use of machine learning algorithms, namely random forest, and support vector machine to extract the root causes from a large accident database in a steel plant, to support ‘root causes analysis for problem solving’. Mayr et al. (2018) and Valamede and Akkari (2020) propose the adoption of data analytics for the 4.0 version of Value Stream Map (VSM), Total Productive Maintenance (TPM), Just-In-Time (JIT), Kanban and poka-yoke. They propose the use of machine learning techniques for (i) VSM 4.0 for automatically generating and validating target states, supporting ‘value identification’; (ii) TPM 4.0 to understand the correlation between condition parameters and probability of defaults and ‘total productive maintenance'; and (iii) poka-yoke 4.0 for the automatic adjustment of machines to irregularities, supporting 'total quality management'.

In conclusion, the reviewed literature shows scholarly interest in adopting DS techniques and tools in LP practices. However, the implementation of DS techniques and tools in LP is still under-investigated, limited to a narrow set of DS techniques and tools and to the 'production planning and control' LP bundle.

3. Methodology

This paper tackles the RQs on the implementation of DS techniques and tools in the PDCA cycle stages related to production, process, inventory, workforce, supplier and customer. The analysis is conducted using the keyword co-occurrence network, a bibliometric analysis tool developed by Waltman, van Eck, and Noyons (2010) and recently applied by different studies (e.g. Strozzi et al. 2017; Cancino et al. 2017; Merigò et al. 2018; Ciano et al. 2019; Xu et al. 2020). This method is based on an approach that combines clusterisation and mapping of bibliometric networks. The modularity-based clustering of VOSViewer is a weighted and parameterised variant of the clustering algorithm developed by Newman and Girvan (2004) to detect communities (clusters) in a network that also considers modularity, a measure of the quality of the clusters structures (Waltman, van Eck, and Noyons 2010). Such analyses are performed by the VoSviewer software (http://www.vosviewer.com/), which gives the keyword co-occurrence network map and the list of the keyword clusters items as results. In the keyword co-occurrence network map, the nodes represent the keywords and the links between nodes represent the occurrence of the two connected keywords in more than a defined amount of papers among the considered ones. The clusters contain keywords that are used more often together, and hence, they refer to the same specific research area. Therefore, the analysis of the obtained clusters can show if the literature has some defined combinations of activities, techniques, and tools of DS in the modelling, management, monitoring, and improvement of production, process, inventory, workforce, supplier, and customer.

The advantage of using this type of literature review is that the literature can be analysed on the basis of a classification and representation of the topics as a result of the quantitative tools implemented by the software, and thus more objective than traditional content-based literature reviews (Strozzi et al. 2017; Kim, Colicchia, and Menachof 2016; Kajikawa et al. 2007).

This method is applied to a data set consisting of the properties and keywords of papers identified through a search in SCOPUS, the largest citation and abstract database of peer-reviewed scientific literature[1] to date (Meester, Steiginga, and Ross 2017; Montoya et al. 2018; Dan et al. 2020; Ribeiro, Fernandes, and Lopes 2020). The search was performed using the TITLE-ABS-KEY field, where ABS is a contraction for abstract, and the KEY field includes AUTHKEY (author keywords) and different kinds of indexed keywords. Author keywords are keywords chosen by the authors themselves to describe the specific content of their work. Indexed keywords are vocabulary and thesaurus terms supplied by a publication to embrace all the content characteristics more broadly and comprehensively (Scopus: Access and use Support Center n.d.). In addition to searching the words in the title, the use of the ABS field can include papers not only exclusively devoted to the topics, but also whose abstract contains a reference to them. Therefore, using all these fields encompasses a relevant number of works.

Regarding DS, the search is completed with the terms ‘big data’, ‘data anal*’ (with the possible ends in -ysis or -ytics) and ‘data’ combined with all the activities of DS identified in Section 2.2. Regarding the fields of implementation, the search involves the

[1] https://www.elsevier.com/__data/assets/pdf_file/0003/966117/Power-of-Scopus_Inforg_Scopus-powers-RI.pdf

PDCA stages combined with the concepts of production, process, inventory, workforce, supplier, and customer related to LP (Shah and Ward 2007; Bai et al. 2019). In May 2020, this search yielded 1157 papers.

The review of the extracted papers helps to understand the subjects identified by the clusters. The clusters are all named by their content to ease their reading and comprehension. As the use of DS techniques and tools in LP practices is meant to improvement, due to its recognized role (Linderman et al. 2003; Jones, Parast and Adams 2010; Watson and DeYong 2010; Dennis and Shook 2007; Liker and Morgan 2011; Marodin and Saurin 2013), and as already done by previous literature reviews (e.g. Chiarini 2011), the subjects are classified considering the PDCA cycle as a framework, and linked to the DS activities identified in Section 2.2.

The main steps of the methodology and the search string explained above are summarised in Figure 1.

Please insert here Figure 1

Once the clusters are analysed, this study searches for the links of the results, obtained by looking at general LP tenets (production, process, inventory, workforce, supplier and customer), to the specific LP practices identified in Section 2.1, identifying the DS techniques and tools that support LP practices and conceptualising a model for their applications following the PDCA cycle.

4. Results

The obtained author keyword co-occurrence network comprises 127 keywords, divided in 15 clusters and ordered by descending number of keywords, as depicted by Figure 2.

Please insert here Figure 2

The clusters are analysed and classified according to PDCA cycle stages, and their content is analysed and linked to the DS activities. The results obtained are summarised in Table 3 and presented in detail in the following subsections.

Please insert here Table 3

4.1 Plan

According to Edwards Deming, '[i]f you cannot describe what you are doing as a process, you do not know what you are doing'. The modelling of business processes, that is the complete description of behavioural aspects of a system to define formal requirements or early design level (Petrasch and Hentschkle 2016), represents the initial stage of any implementation activity. The DS activities from 'data gathering' to 'data modelling and analytics' involve specific techniques and tools. They are exploited with different aims, from the general modelling of business processes to the modelling of business process automation and the development of process models for analytics and process mining.

To gather the data needed to model a process, event logs and Open Platform Communications Unified Architecture (OPC/UA) communication protocol are used (Conforti et al. 2017; Cavalieri, Salafia, and Scroppo 2019; van der Aalst 2018). Log automaton, Petri nets and data flow matrix-based approaches deal with detection of data anomalies, while statistical analysis aims to reduce data dimensionality and perform 'data preparation' activities (Guzikowski et al. 2010; Conforti et al. 2017; Chadli et al. 2018). Ontology, NoSQL databases and relational data modelling support 'data representation', while Extensible Stylesheet Language Transformations enables 'data transformation' (Gruber 2009; Meyer et al. 2013; Krenczyk and Jagodzinski 2015; Meyer et al. 2015; Hassani and Ghannouchi 2017). Feature extraction, clustering, decision trees, association rules, and regression functions are applied to explore data regarding the model to improve its description, its cases, and relations (Kluza et al. 2013; Van der Aalst 2011). MapReduce, supported by Hadoop, is adopted to scale the 'data computing' over the cloud and enable process mining (Sturm, Fichtner, and Schönig 2019). Last, Business Process Model and Notation (BPMN) and text mining contribute to 'data modelling and analytics' adding a process perspective to data and inferring relationships among them (Ligęza 2011; Revina 2019). No specific tool or technique performs 'data visualization and presentation'; however, BPMN is known as a visual tool for business process modelling (Ligęza 2011).

4.2 Do

Reviewed works on DS in the management of production, process, inventory, workforce, supplier and customer involve several DS activities and mainly deal with four topics: knowledge management, smart manufacturing, event-driven process management and data-driven supply chain management.

The management activities are supported by tools to perform 'data gathering', such as the internet of things (IoT), Global Positioning System (GPS), Radio Frequency Identification (RFId) chips, social media, weekly Google index and minutes per viewer of YouTube videos (Fang, Yang and Zong 2018; Gaikwad et al. 2020; He et al. 2017; Heath et al. 2015; Papanagnou and Matthews-Amune 2018; Souza 2014). The text mining algorithms support transformation of the textual data gathered into categories (Walha, Ghozzi and Gargouri 2019). To understand and predict the condition of a system, several techniques and tools perform 'data modelling and analytics', such as simulation, machine learning and deep learning algorithms, hierarchical and k-means clustering, long-short term memory neural network, linear discriminant analysis, image recognition and topic analytics carried out by IBM Alexandria (Heath et al. 2015; Kameswari and Babu 2016; Krumeich, Werth and Loos 2016; Khurana and Kumar 2017; Mehdiyev et al. 2017; Zhuang, Liu and Xiong 2018; Chen et al. 2019; Liu, Miao and Lin 2019; Moreno et al. 2019; Barrad, Gagnon and Valverde 2020; Gaikwad et al. 2020). Management activities are also supported by ‘data visualization and representation’ techniques and tools, such as charts, graphs, tables, maps, data tip, data brushing, dynamic query, and dashboard interfaces (Heath et al. 2015; Husain et al. 2016; Woo et al. 2019).

4.3 Check

Process monitoring techniques have evolved over the last few years, thanks to the recent technological innovations brought by the fourth industrial revolution. Different studies discuss how the DS activities can support the control of the processes, focusing on mainly four topics: (i) the role of the industrial IoT for process monitoring; (ii) the new-generation techniques of multivariate statistical process monitoring; (iii) making big data effective for process monitoring; and (iv) employing process monitoring for quality controls.

Industrial IoT is characterised by a huge number of innovative sensors, usually much cheaper and smaller than the traditional sensors. It can support big ‘data gathering’ and advance the traditional statistical process control (He et al. 2017). To make big data analytics effective, literature studies propose methods to pre-process and handle the data collected to detect outliers, manage missing data and adjust the ranges of independent variables to a common regime (He and Wang 2018; Zhu et al. 2018). Machine learning techniques such as self-organising map and support vector machine have been proposed for ‘data exploration’, perform multiple class classification and ‘data visualization’ when dealing with nonlinear datasets (Tang and Yan 2017; Yan and Yan 2019). Yao and Ge (2018) propose a data computing framework with large datasets based on MapReduce. Finally, among the ‘data analytics’ techniques for process monitoring, a feature-based monitoring statistics pattern analysis emerges as a new methodology in different fields (He, Wang and Shah 2019), as well as predictive modelling algorithms based on machine learning algorithms, statistics and optimisation (Aldrich and Auret 2013; Escobar et al. 2018a; Escobar et al. 2018b; Escobar and Morales-Menendez 2018; Rubab, Taqvi, and Hassan 2019). A specific application of predictive modelling paradigm has been identified in the Process Monitoring for Quality (PMQ), a big data-driven quality control aimed to detect defects through binary classification (Escobar et al. 2018a; Escobar et al. 2018b).

4.4 Act

The DS activities represent an effective decision support for process improvement. Several studies introduce different DS techniques and tools for data-driven decision-making in production environments and show the importance of cloud computing to make big data effective in business process improvement.

Among the DS activities, 'data computing' and, in particular, cloud computing enable quick and efficient data mining and is considered important for making big data effective in business process improvement, as demonstrated empirically by Wang and Zhao (2016). Additionally, data mining techniques are very useful when applied with continuous improvement and quality management. In particular, the most used approach is CRISP-DM (Cross Industry Standard Process for Data Mining) combined with lean six sigma DMAIC (Design, Measure, Analyse, Improve and Control) cycle, which uses modelling techniques and algorithms to improve the quality management activities (Schäfer et al. 2018; Zwetsloot et al. 2018). 'Data representation and transformation' techniques proposed in LSS projects are text mining and video mining (Gupta, Modgil, and Gunasekaran 2020; Ashton, Evangelopoulos, and Prybutok 2015; Yang et al. 2014). 'Data exploration' in LSS projects driven by data mining is possible through the application of data visualisation techniques such as boxplots and histogram (Zwetsloot et al. 2018). While 'data modelling and analytics' techniques proposed for process improvement in quality management are machine learning (Schwenzfeier and Gruhn 2018; Zwetsloot et al. 2018; Ferreira et al. 2015), deep learning (Metzger, Franke, and Jansen 2019; Patil and Thiagarajan 2019; Schäfer et al. 2018) and process mining (Gupta, Modgil, and Gunasekaran 2020). Finally, new algorithms have been recently proposed for 'data visualization' to provide up-to-date and real-time information regarding production processes to a group of manufacturers collaborating on production activities (Shamsuzzoha et al. 2017; Jain, Shao, and Shin 2017).

5. Discussion

5.1 Linking DS tools and techniques and LP practices

The results presented in the previous section are synthetised in Figure 3 which shows what DS techniques and tools can be used to perform the plan, do, check, and act stages of process improvement, i.e. PDCA cycle.

Please insert here Figure 3

DS techniques and tools lie in seven activities that characterise the implementation of DS: (i) data gathering, (ii) data preparation, (iii) data representation and transformation, (iv) data exploration, (v) data computing, (vi) data modelling and analytics and (vii) data visualisation and representation. These seven activities represent a set of sequential fundamental steps that are required to successfully transform massive 'raw data' into meaningful knowledge.

DS techniques and tools help to achieve the purpose of the activities, which are: (i) to collect data, (ii) to check data quality, (iii) to transform and restructure data, (iv) to identify data features, (v) to handle data analysis, (vi) to extract knowledge from data, and (vii) to develop a visual that supports the knowledge sharing.

In the 'Plan' stage (1), the application of DS techniques and tools, such as event logs, statistical analysis, ontology, feature extraction, MapReduce, and text mining, is focused on processes modelling to identify, understand and analyse any areas for improvement, through the process flows (a) and/or the event logs (d), and to identify subsequent implementations of data exchange and process automation (b) and/or for data analytics (c).

In the ‘Do’ stage (2), the application of DS techniques and tool, such as social media, text mining, simulation, and Banana dashboard, is focused on: enabling a better knowledge management using data from additional sources, such as social media, and transforming them into useful information (e); activating specific actions on process management based on predictions obtained from the analysis of process events (g); using the industrial IoT to support choices in production management (f); using data from the actions of customers and suppliers to make decisions guided by objective data in supply chain management (h).

In the ‘Check’ stage (3), the application of DS techniques and tools, such as IIoT, neural networks, machine learning, MapReduce, and machine learning, is focused on: collecting and managing data/big data (k); conducting statistical process control analyses to monitor processes and identify/predict potential problems to be solved (i, j); employing specific applications of statistical process control in the quality management process, using collected data to identify product/process defects (l).

In the ‘Act’ stage (4), the application of DS techniques and tools, such as data pre-processing, text mining, statistics visual analytics, cloud computing, process mining, and dashboard portals, is focused on: providing remedial actions to the problems reported by the check phase, based on a data-driven decision-making guided by a real-time visualisation of the data and forecasting techniques that anticipate any problems (m); speeding up the analysis of large amounts of data to apply process improvements in more competitive times (n); employing specific applications in the field of quality management to improve processes through the application of data mining techniques (o).

Linking these findings to the LP practices definitions provided by the literature, Table 4 was designed to conceptualise these links. In this way, the study answered to the RQs, as discussed in the following sections.

Please insert here Table 4

5.2. Answering RQ1: What DS techniques and tools can support LP practices?

The first aim of this study is to address the gap in the literature related to the link between DS and LP practices, specifically the identification of DS techniques and tools that can support the implementation and boost the results of LP practices.

The result of the bibliometric analysis and its comparison with the LP practices definitions provided by the literature show that the link between DS techniques and tools and LP practices exists. In particular, the findings of this study agree with previous research on the use of DS in support of LP practices that considered the ‘production planning and control’ LP bundle (e.g. Stojanovic et al. 2016; Deuse et al. 2018; Kutschenreiter-Praszkiewicz 2018; Mayr et al. 2018; Abd Rahman et al. 2020; Ito et al. 2020; Sarkar et al. 2020; Shahin et al. 2020; Valamede and Akkari 2020). However, differently from previous research, the results of this research demonstrate also that the application of the identified set of techniques and tools is not limited to the ‘production planning and control’ bundle but can potentially support all the LP bundles. In particular, the findings show what specific DS tools and techniques can support each LP practice.

For example, considering the ‘supplier’ LP bundle, ‘data gathering’ tools (e.g. IoT, GPS, RFId), related to ‘data modelling and analytics’ techniques (e.g. machine learning and deep learning), and ‘data visualization and presentation’ tools (e.g. Dashboards running on an Apache web server), can support generating the content of the reporting on performance for the ‘feedback to suppliers’, addressing the definition of the product parameters for the ‘JIT delivery by supplier’, efficiently and effectively sharing the generated knowledge for the ‘supplier involvement in design’, and audit the status of lean development for the ‘lean supplier development’.

Whereas, for example, considering the ‘production planning and control’ bundle, the use of ‘data gathering’ tools (e.g. RFId, Pad, indoor GPS) and ‘data modelling and analytics’ techniques (e.g. machine learning and deep learning algorithms, image recognition), can support translating elementary motions into activities and setting manual task time standard for ‘setup reduction’ (as reported by Kutschenreiter-Praszkiewicz 2018), fault detection for ‘total productive maintenance’, identification of variability for ‘total quality management’, and the identification of the thresholds required for real-time processing and mine process anomalies (as reported by Stojanovic and Milenovic, 2018).

5.3. Answering RQ2: How should DS techniques and tools be applied to support LP practices?

The second aim of this study is to define a model to guide the application of DS techniques and tools to support LP practices. The results obtained in this study show that DS techniques and tools, and the sequential DS activities to which they belong, can support LP practices based on the typical lean implementation stages of the PDCA methodology. The findings in Figure 3 and Table 4, according to previous literature recognising PDCA as an improvement framework (Linderman et al. 2003; Jones, Parast and Adams 2010; Watson and DeYong 2010; Dennis and Shook 2007; Liker and Morgan 2011; Marodin and Saurin 2013), identify the PDCA cycle as a framework which makes it possible for all the DS activities to support the whole set of LP practices. Thus, the authors developed a model, depicted in Figure 3, with the aim to guide both selection and application of DS techniques and tools to support LP. In particular, the model outlines how the DS tools and techniques should be adopted according to a set of pursued objectives of each stage of the PDCA cycle, following the seven DS activities.

For example, ‘JIT delivery to supplier’ activity can be supported by DS techniques and tools in four stages: (i) plan: modelling processes and identifying, understanding and analysing the involvement of suppliers for process data exchange; (ii) do: addressing the definition of the product parameters; (iii) check: defining the status of quantities, quality and time of the deliveries; (iv) act: activating a collaborative process management between partners.

6. Conclusions

Starting from a relevant gap in the literature, this study links DS to LP through a bibliometric analysis of the literature on the implementations of DS techniques and tools applied to the key tenets of LP (production, process, inventory, workforce, supplier and customer). The results help develop a model that supports this link grounded on the PDCA model.

This study contributes to the topics of LP and DS in various significant ways. Theoretically, first, it unveils a research gap by highlighting the scarcity of research linking DS tools and techniques to LP practices. Third, based on the reviewed literature on DS techniques and tools applied to the key LP tenets in each stage of the PDCA cycle, this work develops a model that identifies what DS techniques and tools can support each LP practice and how to achieve this goal (Figure 3 and Table 4). Finally, the results of this study offer a base for future research that could explore each link and conduct empirical research to validate the proposed implementation model.

From a practical viewpoint, the described links can inspire practitioners about exploiting the value of data to facilitate and boost the traditional LP practices. The developed model can serve as a guide for using the DS techniques and tools to support LP practices, according to the specific stages of the PDCA cycle. Moreover, practitioners can understand and anticipate the requirements for undertaking specific improvement decisions in this context. The reference to specific DS tools and techniques in the area they want to improve can suggest searching for the most effective software or platforms and investing in training for their implementation or in supporting technologies.

Despite the contributions listed, this study has some limitations. Regarding the methodology adopted (Figure 1), first, the definition of the links is based on the analysis of keyword co-occurrence networks generated according to the analysed dataset derived only from Scopus, that, although quite comprehensive, includes only a fraction of scientific publications (Strozzi et al. 2017). Second, the method does not involve an empirical validation step. Therefore, future research could work towards overcoming these limitations, empirically addressing the refinement and validation of this study results.

Moreover, regarding the model, it is oriented to companies that have been implementing LP practices, know DS techniques and tools, and have the competences to successfully implement the model. Accordingly, future research could consider the critical success factors that a company needs to accomplish the implementation of this model. Last, the proposed model cannot be extended to a Six Sigma-based production system, based on DMAIC cycle, in which specific tools are used in different steps (Choo, Linderman and Schroeder 2007). Thus, extending the proposed model for exploiting the value of data to facilitate and boost the specific tools involved by DMAIC can represent a future research direction.

Declaration of Interest: none.

References

Abd Rahman, M.S., Mohamad, E. and Rahman, A.A., 2020. Enhancement of overall equipment effectiveness (OEE) data by using simulation as decision making tools for line balancing. Indonesian Journal of Electrical Engineering and Computer Science, 18(2), pp.1040-1047.

Addo-Tenkorang, R. and Helo, P.T., 2016. Big data applications in operations/supply-chain management: A literature review. Computers & Industrial Engineering, 101, pp.528-543.

Agarwal, R. and Dhar, V., 2014. Big Data, Data Science, and Analytics: The Opportunity and Challenge for IS Research. Information Systems Research, 25(3), pp.443-448.

Agostini, Lara, and Roberto Filippini. 2019. Organizational and Managerial Challenges in the Path toward Industry 4.0. European Journal of Innovation Management. Emerald Publishing Limited.

Akter, S., Wamba, S.F., Gunasekaran, A., Dubey, R. and Childe, S.J., 2016. How to improve firm performance using big data analytics capability and business strategy alignment?. International Journal of Production Economics, 182, pp.113-131.

Aldrich, C. and Auret, L., 2013. Unsupervised process monitoring and fault diagnosis with machine learning methods. London: Springer.

Alkhoraif, A., Rashid, H. and McLaughlin, P., 2019. Lean implementation in small and medium enterprises: Literature review. Operations Research Perspectives, 6, p.100089.

Antony, J. and Sony, M., 2019. An empirical study into the limitations and emerging trends of Six Sigma in manufacturing and service organisations. International Journal of Quality & Reliability Management.

Antony, J., Gupta, S. and Gijo, E.V., 2018. Ten commandments of Lean Six Sigma: a practitioners' perspective. International Journal of Productivity and Performance Management.

Ashton, Triss, Nicholas Evangelopoulos, and Victor R Prybutok. 2015. "Quantitative Quality Control from Qualitative Data: Control Charts with Latent Semantic Analysis." Quality & Quantity 49 (3). Springer: 1081–1099.

Baesens, B., 2014. Analytics in a big data world: The essential guide to data science and its applications. John Wiley & Sons.

Bai, C., Satir, A., and Sarkis, J. 2019. Investing in lean manufacturing practices: an environmental and operational perspective. International Journal of Production Research, 57(4), 1037-1051.

Bailer, A.J. and Fisher, N.I., 2020. Discussion of "A review of data science in business and industry and a future view". Applied Stochastic Models in Business and Industry, 36(1), pp.20-22.

Ballé, F. and Ballé, M., 2005. Lean development. Business Strategy Review, 16(3), pp.17-22.

Barrad, S., Gagnon, S. and Valverde, R., 2020. An Analytics Architecture for Procurement. International Journal of Information Technologies and Systems Approach (IJITSA), 13(2), pp.73-98.

Baryannis, George, Sahar Validi, Samir Dani, and Grigoris Antoniou. 2019. "Supply Chain Risk Management and Artificial Intelligence: State of the Art and Future Research Directions." International Journal of Production Research 57 (7). Taylor & Francis: 2179–2202.

Bazrkar, A. and Iranzadeh, S., 2017. Prioritization of Lean Six Sigma Improvement Projects using Data Envelopment Analysis Cross Efficiency Model. Calitatea, 18(157), p.72.

Belekoukias, I., Garza-Reyes, J.A. and Kumar, V., 2014. The impact of lean methods and tools on the operational performance of manufacturing organisations. International Journal of production research, 52(18), pp.5346-5366.

Belhadi, A., Kamble, S.S., Zkik, K., Cherrafi, A. and Touriki, F.E., 2020. The integrated effect of Big Data Analytics, Lean Six Sigma and Green Manufacturing on the environmental performance of manufacturing companies: The case of North Africa. Journal of Cleaner Production, 252, p.119903.

Bevilacqua, M., Ciarapica, F.E. and De Sanctis, I., 2017. Lean practices implementation and their relationships with operational responsiveness and company performance: an Italian study. International Journal of Production Research, 55(3), pp.769-794.

Bhamu, J. and Sangwan, K.S., 2014. Lean manufacturing: literature review and research issues. International Journal of Operations & Production Management, 34 (7), pp. 876-940.

Bittencourt, V L, Anabela Carvalho Alves, and Celina Pinto Leão. 2021. Industry 4.0 Triggered by Lean Thinking: Insights from a Systematic Literature Review. International Journal of Production Research 59 (5). Taylor & Francis: 1496–1510.

Bortolotti, T., S. Boscari, and P. Danese. 2015. “Successful Lean Implementation: Organizational Culture and Soft Lean Practices.” International Journal of Production Economics 160: 182–201.

Buer, S. V., Strandhagen, J. O., and Chan, F. T. 2018. The link between Industry 4.0 and lean manufacturing: mapping current research and establishing a research agenda. International Journal of Production Research, 56(8), 2924-2940.

Buer, S.V., Semini, M., Strandhagen, J.O. and Sgarbossa, F., 2020. The complementary effect of lean manufacturing and digitalisation on operational performance. International Journal of Production Research, pp.1-17.

Cancino, C., Merigó, J.M., Coronado, F., Dessouky, Y. and Dessouky, M., 2017. Forty years of Computers & Industrial Engineering: A bibliometric analysis. Computers & Industrial Engineering, 113, pp.614-629.

Cao, L., 2017. Data science: a comprehensive overview. ACM Computing Surveys (CSUR), 50(3), pp.1-42.

Cavalieri, S., Salafia, M.G. and Scroppo, M.S., 2019. Integrating OPC UA with web technologies to enhance interoperability. Computer Standards & Interfaces, 61, pp.45-64.

Chadli, N., Kabbaj, M. I., & Bakkoury, Z. 2018, October. Detection of Dataflow Anomalies in Business Process An Overview of Modeling Approaches. In Proceedings of the 12th International Conference on Intelligent Systems: Theories and Applications (pp. 1-6).

Chawla, G., Bamal, S., & Khatana, R. (2018). Big Data Analytics for Data Visualization: Review of Techniques'. International Journal of Computer Applications, 182(21), 37-40.

Chen, C., Wang, Y., Huang, G., & Xiong, H. 2019, December. Hierarchical Demand Forecasting for Factory Production of Perishable Goods. In 2019 IEEE International Conference on Big Data (Big Data) (pp. 188-193). IEEE.

Cheng, Y., Chen, K., Sun, H., Zhang, Y., & Tao, F. 2018. Data and knowledge mining with big data towards smart production. Journal of Industrial Information Integration, 9, 1-13.

Chiarini, A. 2011. Japanese total quality control, TQM, Deming's system of profound knowledge, BPR, Lean and Six Sigma. International Journal of Lean Six Sigma.

Chinosi, M., and Trombetta, A. 2012. BPMN: An introduction to the standard. Computer Standards & Interfaces, 34(1), 124-134.

Choo, A.S., Linderman, K.W. and Schroeder, R.G., 2007. Method and context perspectives on learning and knowledge creation in quality management. Journal of operations management, 25(4), pp.918-931.

Ciano, M. P., Pozzi, R., Rossi, T., and Strozzi, F. 2019. How IJPR has addressed ‘lean’: a literature review using bibliometric tools. International Journal of Production Research, 1-34.

Ciano, M.P., Dallasega, P., Orzes, G. and Rossi, T., 2020. One-to-one relationships between Industry 4.0 technologies and Lean Production techniques: a multiple case study. International Journal of Production Research, pp.1-25.

Concolato, C.E. and Chen, L.M., 2017. Data science: A new paradigm in the age of big-data science and analytics. New Mathematics and Natural Computation, 13(02), pp.119-143.

Conforti, V., Bulgarelli, A., Fioretti, V., Gianotti, F., Malaguti, G., Trifoglio, M., Antonelli, L.A., Gallozzi, S., Lombardi, S., Lucarelli, F. and Mastropietro, M., 2017. Procedures of software integration test and release for ASTRI SST-2M prototype proposed for the Cherenkov Telescope Array. In Proc. ICALEPS.

Cooper, M.C., Lambert, D.M. and Pagh, J.D., 1997. Supply chain management: more than a new name for logistics. The international journal of logistics management, 8(1), pp.1-14.

Cua, K. O., K. E. McKone, and R. G. Schroeder. 2001. "Relationships Between Implementation of TQM, JIT, and TPM and Manufacturing Performance." Journal of Operations Management 19 (6): 675–694.

Dahlgaard-Park, S.M. and Pettersen, J., 2009. Defining lean production: some conceptual and practical issues. The TQM journal.

Dan, Z., Jiapeng, X., Yizhu, Z., Jing, W., Siyu, H. and Xiao, Z., 2020. Study on sustainable urbanization literature based on Web of Science, scopus, and China national knowledge infrastructure: A scientometric analysis in CiteSpace. Journal of cleaner production, p.121537.

Davenport, T.H. and Patil, D.J., 2012. Data scientist. Harvard business review, 90(5), pp.70-76.

Demirkan, H., and Delen, D. 2013. Leveraging the capabilities of service-oriented decision support systems: Putting analytics and big data in cloud. Decision Support Systems, 55(1), 412-421.

Dennis, P. and Shook, J. 2007. Lean Production Simplified: A Plain Language Guide to the World's Most Powerful Production Systems, Productivity Press, New York, NY.

Detty, R.B. and Yingling, J.C., 2000. Quantifying benefits of conversion to lean manufacturing with discrete event simulation: a case study. International Journal of Production Research, 38(2), pp.429-445.

Deuse, J., Heuser, C., Konrad, B., Lenze, D., Maschek, T., Wiegand, M. and Willats, P., 2018. Pushing the limits of lean thinking–Design and management of complex production systems. In Closing the gap between practice and research in industrial engineering (pp. 335-342). Springer, Cham.

Diop, Mouhamed, et al. "Prior management of temporal data quality in a data mining process: an implementation architecture." Procedia computer science 148. 2019. 273-282.

Dogan, O. and Gurcan, O.F., 2018, July. Data perspective of Lean Six Sigma in industry 4.0 Era: a guide to improve quality. In Proceedings of the international conference on industrial engineering and operations management Paris.

Donoho, David. 2017. "50 Years of Data Science." Journal of Computational and Graphical Statistics 26 (4). Taylor & Francis: 745–766.

Doolen, T. L., and Hacker, M. E. 2005. A review of lean assessment in organizations: an exploratory study of lean practices by electronics manufacturers. Journal of Manufacturing systems, 24(1), 55-67.

Escobar, C. A., and Morales-Menendez, R. 2019. Process-monitoring-for-quality—A robust model selection criterion for the logistic regression algorithm. Manufacturing Letters, 22, 6-10.

Escobar, Carlos A, Jeffrey A Abell, Marcela Hernández-de-Menéndez, and Ruben Morales-Menendez. 2018. "Process-Monitoring-for-Quality—Big Models." Procedia Manufacturing 26. Elsevier: 1167–1179.

Escobar, Carlos A, Michael A Wincek, Debejyo Chakraborty, and Ruben Morales-Menendez. 2018. "Process-Monitoring-for-Quality—Applications." Manufacturing Letters 16. Elsevier: 14–17.

Eybers, S. and Mayet, R., 2020, May. From Data to Insight: A Case Study on Data Analytics in the Furniture Manufacturing Industry. In International Conference on Integrated Science (pp. 392-405). Springer, Cham.

Fang, P., Jiang, Y. and Zhong, R.Y., 2018, December. Real-time monitoring of workshop status based on internet of things. In Proceedings of the 2018 48th IEEE International Conference on Computers and Industrial Engineering (CIE 48), Auckland, New Zealand (pp. 2-5).

Farooqui, A., Bengtsson, K., Falkman, P. and Fabian, M., 2020. Towards data-driven approaches in manufacturing: an architecture to collect sequences of operations. International Journal of Production Research, pp.1-17.

Ferreira, Filipe, Ahm Shamsuzzoha, Americo Azevedo, and Petri Helo. 2015. "Virtual Enterprise Process Monitoring: An Approach towards Predictive Industrial Maintenance." In Progress in Systems Engineering, 285–291. Springer.

Gaikwad, A., Yavari, R., Montazeri, M., Cole, K., Bian, L., and Rao, P. 2020. Toward the digital twin of additive manufacturing: Integrating thermal simulations, sensing, and analytics to detect process faults. IISE Transactions, 1-14.

Gamal Aboelmaged, M. 2010. "Six Sigma Quality: a Structured Review and Implications for Future Research." International Journal of Quality & Reliability Management 27 (3): 268–317.

Garre, A., Ruiz, M. C., and Hontoria, E. 2020. Application of Machine Learning to support production planning of a food industry in the context of waste generation under uncertainty. Operations Research Perspectives, 100147.

Gorodov, E. Y. E., and Gubarev, V. V. E. (2013). Analytical review of data visualization methods in application to big data. Journal of Electrical and Computer Engineering, 2013.

Greasley, A., and Edwards, J. S. (2019). Enhancing discrete-event simulation with big data analytics: A review. Journal of the Operational Research Society, 1-21.

Gruber, T. 2009. Ontology. In: Ling L, Tamer Özsu M (eds) The encyclopedia of database systems. Springer, pp 1963–1965.

Gupta, S., Modgil, S. and Gunasekaran, A., 2020. Big data in lean six sigma: a review and further research directions. International Journal of Production Research, 58(3), pp.947-969.

Guzikowski, S., Bergum, J., Cassidy, M., Dong, L., Lai, C., Patel, B., Randazzo, M., Razler, T., Reiff, E., Rosso, V., Rubin, E., Tabora, J., Thornton, J., Xu, Y. 2010. Applications of model-based quality by design for reaction engineering. AIChE Annual Meeting, Salt Lake City, Utah.

Hahn, G.J., 2020. Industry 4.0: a supply chain innovation perspective. International Journal of Production Research, 58(5), pp.1425-1441.

Hajirahimova, M., and Ismayilova, M., 2018. Big data visualization: Existing approaches and problems. Problems of Information Technology, 9, 72-83.

Hashemian, H.M., 2010. State-of-the-art predictive maintenance techniques. IEEE Transactions on Instrumentation and measurement, 60(1), pp.226-236.

Hassani, A., and Ghannouchi, S. A. 2017, December. Exploring the Integration of Business Process with Nosql Databases in the Context of BPM. In International Conference on Intelligent Systems Design and Applications (pp. 771-784). Springer, Cham.

Hazen, B.T., Boone, C.A., Ezell, J.D. and Jones-Farmer, L.A., 2014. Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. International Journal of Production Economics, 154, pp.72-80.

Hazen, B.T., Skipper, J.B., Boone, C.A. and Hill, R.R., 2018. Back in business: Operations research in support of big data analytics for operations and supply chain management.

Annals of Operations Research, 270(1-2), pp.201-211.

He, Q Peter, Jin Wang, Devarshi Shah, and Nader Vahdat. 2017. “Statistical Process Monitoring for IoT-Enabled Cybermanufacturing: Opportunities and Challenges.” IFAC-PapersOnLine 50 (1). Elsevier: 14946–14951.

He, Q. P., Wang, J., and Shah, D. (2019). Feature space monitoring for smart manufacturing via statistics pattern analysis. Computers & Chemical Engineering, 126, 321-331.

He, Q.P. and Wang, J., 2018. Statistical process monitoring as a big data analytics tool for smart manufacturing. Journal of Process Control, 67, pp.35-43.

Heath, F.F., Hull, R., Khabiri, E., Riemer, M., Sukaviriya, N. and Vaculín, R., 2015, June. Alexandria: extensible framework for rapid exploration of social media. In 2015 IEEE International Congress on Big Data (pp. 483-490). IEEE.

Heger, J., and El Abdine, M. Z. 2019. Using Data Mining Techniques to Investigate the Correlation between Surface Cracks and Flange Lengths in Deep Drawn Sheet Metals. IFAC-PapersOnLine, 52(13), 851-856.

Husain, W., Koon, L.W., Zhao, L.T. and Aziz, N.T.B.A., 2016, August. A proposed framework for enhancing a supply chain management system to support flood disaster relief operations. In 2016 3rd International Conference on Computer and Information Sciences (ICCOINS) (pp. 25-30). IEEE.

Ito, T., Abd Rahman, M.S., Mohamad, E., Abd Rahman, A.A. and Salleh, M.R., 2020. Internet of things and simulation approach for decision support system in lean manufacturing. Journal of Advanced Mechanical Design, Systems, and Manufacturing, 14(2).

Jagadish, Hosagrahar Visvesva. 2015. “Big Data and Science: Myths and Reality.” Big Data Research 2 (2). Elsevier: 49–52.

Jain, Sanjay, Guodong Shao, and Seung-Jun Shin. 2017. “Manufacturing Data Analytics

Using a Virtual Factory Representation." International Journal of Production Research 55 (18). Taylor & Francis: 5450–5464.

Jasti, N. V. K., and Kodali, R. 2015. Lean production: literature review and trends. International Journal of Production Research, 53(3), 867-885.

Joe Qin, S., 2003. Statistical process monitoring: basics and beyond. Journal of Chemometrics: A Journal of the Chemometrics Society, 17(8-9), pp.480-502.

Jones, E. C., Parast, M. M., and Adams, S. G. 2010. A framework for effective Six Sigma implementation. Total quality management, 21(4), 415-424.

Kajikawa, Y., J. Ohno, Y. Takeda, K. Matsushima, and H. Komiyama. 2007. "Creating an Academic Landscape of Sustainability Science: An Analysis of the Citation Network." Sustainability Science 2 (2): 221–231.

Kamble, Sachin S, and Angappa Gunasekaran. 2020. "Big Data-Driven Supply Chain Performance Measurement System: A Review and Framework for Implementation." International Journal of Production Research 58 (1). Taylor & Francis: 65–86.

Kameswari, U Surya, and I Ramesh Babu. 2016. "Sensor Data Analysis and Anomaly Detection Using Predictive Analytics for Process Industries." In *2015 IEEE Workshop on Computational Intelligence: Theories, Applications and Future Directions (WCI)*, 1–8. IEEE.

Karlsson, C., and P. Åhlström. 1996. "Assessing Changes Towards Lean Production." International Journal of Operations & Production Management 16 (2): 24–41.

Keivanpour, S. and Kadi, D.A., 2019. The Effect of "Internet of Things" on Aircraft Spare Parts Inventory Management. IFAC-PapersOnLine, 52(13), pp.2343-2347.

Khalid, Z. M., and Zebaree, S. R. (2021). Big Data Analysis for Data Visualization: A Review. International Journal of Science and Business, 5(2), 64-75.

Khurana, M. and Kumar, D., 2017, December. The study of data analytics in inventory management. In 2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions)(ICTUS) (pp. 140-144). IEEE.

Kim, S., C. Colicchia, and D. Menachof. 2016. “Ethical Sourcing: An Analysis of the Literature and Implications for Future Research.” Journal of Business Ethics: 1–20. doi:10.1007/s10551-016-3266-8.

Kingsley, O., Tawil, A. R. H., Naeem, U., Islam, S., and Lamine, E. 2016, December. Using semantic-based approach to manage perspectives of process mining: Application on improving learning process domain data. In 2016 IEEE international conference on big data (big data) (pp. 3529-3538). IEEE.

Kipper, L.M., Furstenau, L.B., Hoppe, D., Frozza, R. and Iepsen, S. 2020. Scopus scientific mapping production in industry 4.0 (2011–2018): a bibliometric analysis, International Journal of Production Research, 58 (6), 1605-1627.

Kluza, K., Baran, M., Bobek, S., and Nalepa, G. J. 2013. Overview of recommendation techniques in business process modeling. In Proceedings of 9th Workshop on Knowledge Engineering and Software Engineering (KESE9), pp. 46-57.

Köksal, G., Batmaz, İ. and Testik, M.C., 2011. A review of data mining applications for quality improvement in manufacturing industry. Expert systems with Applications, 38(10), pp.13448-13467.

Krafcik, J.F., 1988. Triumph of the lean production system. Sloan management review, 30(1), pp.41-52.

Krenczyk, D., and Jagodzinski, M. 2015. ERP, APS and simulation systems integration to support production planning and scheduling. In 10th International Conference on Soft Computing Models in Industrial and Environmental Applications (pp. 451-461). Springer, Cham.

Krumeich, Julian, Dirk Werth, and Peter Loos. 2016. “Prescriptive Control of Business

Processes.” Business & Information Systems Engineering 58 (4). Springer: 261–280.

Kumar, Sandeep, Bhushan S. Purohit, Vikas Manjrekar, Vivek Singh, and Bhupesh Kumar Lad. 2018. “Investigating the Value of Integrated Operations Planning: A Case-Based Approach From Automotive Industry.” International Journal of Production Research 56 (22): 6971–6992.

Kuonen, D. and Loison, B., 2019. Production processes of official statistics and analytics processes augmented by trusted smart statistics: Friends or foes?. Statistical Journal of the IAOS, 35(4), pp.615-622.

Kusiak, A., 2018. Smart manufacturing. International Journal of Production Research, 56(1-2), pp.508-517.

Kutschenreiter-Praszkiewicz, I. (2018). Machine learning in SMED. Journal of Machine Engineering, 18.

Ligęza, A. 2011. BPMN–a logical model and property analysis. Decision Making in Manufacturing and Services, 5(1-2), 57-67.

Liker, J. K., and Morgan, J. 2011. Lean product development as a system: a case study of body and stamping development at Ford. Engineering Management Journal, 23(1), 16-28.

Liker, J.K., 2004. Toyota way: 14 management principles from the world's greatest manufacturer. McGraw-Hill Education.

Linderman, K., Schroeder, R.G., Zaheer, S. and Choo, A.S. 2003. “Six Sigma: a goal-theoretic perspective”, Journal of Operations Management, Vol. 3 No. 21, pp. 193-203.

Liu, Q., Miao, K. and Lin, K., 2019, June. Inventory Management of Automobile After-sales Parts Based on Data Mining. In Proceedings of the 2019 3rd High Performance Computing and Cluster Technologies Conference (pp. 195-199).

Lugert, A., Batz, A. and Winkler, H., 2018. Empirical assessment of the future adequacy of value stream mapping in manufacturing industries. Journal of Manufacturing Technology Management, 29(5), pp.886-906.

MacDuffie, J.P., 1995. Human resource bundles and manufacturing performance: Organizational logic and flexible production systems in the world auto industry. Ilr Review, 48(2), pp.197-221.

Mann, D., 2005. Creating a lean culture: tools to sustain lean conversion. New York: Productivity Press.

Marksberry, P., 2012. Investigating “The Way” for Toyota suppliers: A quantitative outlook on Toyota's replicating efforts for supplier development. Benchmarking: An International Journal.

Marodin, G.A. and Saurin, T.A., 2013. Implementing lean production systems: research areas and opportunities for future studies. International Journal of Production Research, 51(22), pp.6663-6680.

Mayr, A., Weigelt, M., Kühl, A., Grimm, S., Erll, A., Potzel, M., and Franke, J. 2018. Lean 4.0-A conceptual conjunction of lean management and Industry 4.0. Procedia Cirp, 72, 622-628.

McIntosh, R. I., Culley, S. J., Mileham, A. R., and Owen, G. W. (2000). A critical evaluation of Shingo's' SMED'(Single Minute Exchange of Die) methodology. International Journal of Production Research, 38(11), 2377-2395.

Meester, W.J., Steiginga, S. and Ross, C.A., 2017. A brief history of Scopus: The world’s largest abstract and citation database of scientific literature. Research analytics: Boosting university productivity and competitiveness through Scientometrics, p.31.

Mehdiyev, Nijat, Andreas Emrich, Björn P Stahmer, Peter Fettke, and Peter Loos. 2017. “IPRODICT-Intelligent Process Prediction Based on Big Data Analytics.” In BPM (Industry Track), 13–24.

Mennens, Robin J P, Roeland Scheepens, and Michel A Westenberg. 2019. “A Stable Graph Layout Algorithm for Processes.” In Computer Graphics Forum, 38:725–737. Wiley Online Library.

Merigó, J.M., Pedrycz, W., Weber, R. and de la Sotta, C., 2018. Fifty years of Information Sciences: A bibliometric overview. Information Sciences, 432, pp.245-268.

Metzger, A., Franke, J. and Jansen, T., 2019. Data-driven deep learning for proactive terminal process management. In BPM (Industry Forum) (pp. 190-201).

Meudt, T., J. Metternich, and E. Abele. 2017. “Value Stream Mapping 4.0: Holistic Examination of Value Stream and Information Logistics in Production.” CIRP Annals – Manufacturing Technology 66 (1): 413–416.

Meyer, A., Pufahl, L., Batoulis, K., Fahland, D., and Weske, M. 2015. Automating data exchange in process choreographies. Information Systems, 53, 296-329.

Meyer, A., Pufahl, L., Fahland, D., and Weske, M. 2013. Modeling and enacting complex data dependencies in business processes. In Business process management (pp. 171-186). Springer, Berlin, Heidelberg.

Möldner, A.K., Garza-Reyes, J.A. and Kumar, V., 2020. Exploring lean manufacturing practices' influence on process innovation performance. Journal of Business Research, 106, pp.233-249.

Montoya, F.G., Alcayde, A., Baños, R. and Manzano-Agugliaro, F., 2018. A fast method for identifying worldwide scientific collaborations using the Scopus database. Telematics and Informatics, 35(1), pp.168-185.

Moreno, R., Pereira, J. C., López, A., Mohammed, A., and Pahlevannejad, P. 2019, November. Time Series Display for Knowledge Discovery on Selective Laser Melting Machines. In International Conference on Intelligent Data Engineering and Automated Learning (pp. 280-290). Springer, Cham.

Mourtzis, D., Papathanasiou, P. and Fotia, S., 2016. Lean rules identification and classification for manufacturing industry. Procedia CIRP, 50, pp.198-203.

Negrão, L.L.L., Godinho Filho, M. and Marodin, G., 2017. Lean practices and their effect on performance: a literature review. Production Planning & Control, 28(1), pp.33-56.

Newman, M.E. and Girvan, M., 2004. Finding and evaluating community structure in networks. Physical review E, 69(2), p.026113.

Noori, B. and Latifi, M., 2018. Development of Six Sigma methodology to improve grinding processes. International journal of lean six sigma.

Osterman, P., 1994. How common is workplace transformation and who adopts it?. ILR Review, 47(2), pp.173-188.

Panizzolo, R. 1998. “Applying the Lessons Learned From 27 Lean Manufacturers.: The Relevance of Relationships Management.” International Journal of Production Economics 55 (3): 223–240.

Papanagnou, C.I. and Matthews-Amune, O., 2018. Coping with demand volatility in retail pharmacies with the aid of big data exploration. Computers & Operations Research, 98, pp.343-354.

Park, S.H., Dahlgaard-Park, S.M. and Kim, D.C., 2020. New Paradigm of Lean Six Sigma in the 4th Industrial Revolution Era. Quality Innovation Prosperity, 24(1), pp.1-16.

Patil, Ranjeet, and Balaji Thiagarajan. 2019. “Design Optimization through Blended Reliability, Six Sigma and Analytics Tools.” In 2019 Annual Reliability and Maintainability Symposium (RAMS), 1–3. IEEE.

Petrasch, R. and Hentschke, R., 2016, July. Process modeling for Industry 4.0 applications: Towards an Industry 4.0 process modeling language and method. In 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE) (pp. 1-5). IEEE.

Pettersen, J. 2009. “Defining Lean Production: Some Conceptual and Practical Issues.” The TQM Journal 21 (2): 127–142.

Pettersen, K. J., Handfield, R. B., and Ragatz, G. L. 2005. Supplier integration into new product development: coordinating product, process and supply chain design. Journal of operations management, 23(3-4), 371-388.

Phuong, N.A. and Guidat, T., 2018, July. Sustainable value stream mapping and technologies of Industry 4.0 in manufacturing process reconfiguration: A case study in an apparel company. In 2018 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI) (pp. 85-90). IEEE.

Provost, F., and Fawcett, T. 2013. “Data Science and Its Relationship to Big Data and Data-Driven Decision-making.” Big Data 1 (1). Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA: 51–59.

Queiroz, M.M., Wamba, S.F., Machado, M.C. and Telles, R., 2020. Smart production systems drivers for business process management improvement. Business Process Management Journal.

Rai, S., 2019. Big data-real time fact-based decision: the next big thing in supply chain. International Journal of Business Performance and Supply Chain Modelling, 10(3), pp.253-265.

Raman, S., Patwa, N., Niranjan, I., Ranjan, U., Moorthy, K. and Mehta, A., 2018. Impact of big data on supply chain management. International Journal of Logistics Research and Applications, 21(6), pp.579-596.

Rejikumar, G., Aswathy Asokan, A. and Sreedharan, V.R., 2020. Impact of data-driven decision-making in Lean Six Sigma: an empirical analysis. Total Quality Management & Business Excellence, 31(3-4), pp.279-296.

Revina, A. 2018, July. Assessing Process Suitability for AI-Based Automation. Research

Idea and Design. In International Conference on Business Information Systems (pp. 697-706). Springer, Cham

Ribeiro, M.I.B., Fernandes, A.J.G. and Lopes, I.M., 2020. Digital marketing: A bibliometric analysis based on the Scopus database scientific publications. In Digital Marketing Strategies and Models for Competitive Business (pp. 52-73). IGI Global.

Rosin, F., Forget, P., Lamouri, S., and Pellerin, R. 2020. Impacts of Industry 4.0 technologies on Lean principles. International Journal of Production Research, 58 (6) 1644-1661.

Rubab, Saddaf, Syed A Taqvi, and Mohd Fadzil Hassan. 2019. “Realizing the Value of Big Data in Process Monitoring and Control: Current Issues and Opportunities.” In International Conference of Reliable Information and Communication Technology, 128–138. Springer.

Sancha, C., Wiengarten, F., Longoni, A. and Pagell, M., 2020. The moderating role of temporary work on the performance of lean manufacturing systems. International Journal of Production Research, 58(14), pp.4285-4305.

Santos, F., Pereira, R. and Vasconcelos, J.B., 2019. Toward robotic process automation implementation: an end-to-end perspective. Business Process Management Journal, 26(2), pp.405-420.

Sarkar, S., Ejaz, N., Kumar, M., and Maiti, J. 2020. Root Cause Analysis of Incidents Using Text Clustering and Classification Algorithms. In Proceedings of ICETIT 2019 (pp. 707-718). Springer, Cham.

Saurin, T. A., G. A. Marodin, and J. L. D. Ribeiro. 2011. “AFramework for Assessing the use of Lean Production Practices in Manufacturing Cells.” International Journal of Production Research 49 (11): 3211–3230.

Schäfer, Franziska, Christian Zeiselmair, Jonas Becker, and Heiner Otten. 2018. “Synthesizing CRISP-DM and Quality Management: A Data Mining Approach for

Production Processes." In 2018 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD), 190–195. IEEE.

Schwenzfeier, Nils, and Volker Gruhn. 2018. "Toward a Practical Process Model for Anomaly Detection Systems." In 2018 IEEE/ACM 1st International Workshop on Software Engineering for Cognitive Services (SE4COG), 41–44. IEEE.

Sezen, B., I. S. Karakadilar, and G. Buyukozkan. 2012. "Proposition of a Model for Measuring Adherence to Lean Practices: Applied to Turkish Automotive Part Suppliers." International Journal of Production Research 50 (14): 3878–3894.

Shah, Devarshi, Jin Wang, and Q Peter He. 2019. "An Internet-of-Things Enabled Smart Manufacturing Testbed." IFAC-PapersOnLine 52 (1). Elsevier: 562–567.

Shah, R., and P. T. Ward. 2003. "Lean Manufacturing: Context, Practice Bundles, and Performance." Journal of Operations Management 21 (2): 129–149.

Shah, R., and P. T. Ward. 2007. "Defining and Developing Measures of Lean Production." Journal of Operations Management 25 (4): 785–805.

Shah, R., Chandrasekaran, A. and Linderman, K., 2008. In pursuit of implementation patterns: the context of Lean and Six Sigma. International Journal of Production Research, 46(23), pp.6679-6699.

Shahin, M., Chen, F.F., Bouzary, H. and Krishnaiyer, K., 2020. Integration of Lean practices and Industry 4.0 technologies: smart manufacturing for next-generation enterprises. The International Journal of Advanced Manufacturing Technology, pp.1-10.

Shamsuzzoha, Ahm, Filipe Ferreira, Americo Azevedo, and Petri Helo. 2017. "Collaborative Smart Process Monitoring within Virtual Factory Environment: An Implementation Issue." International Journal of Computer Integrated Manufacturing 30 (1). Taylor & Francis: 167–181.

Shao, G., Shin, S.J. and Jain, S., 2014, December. Data analytics using simulation for smart manufacturing. In Proceedings of the Winter Simulation Conference 2014 (pp. 2192-2203). IEEE.

Sharma, S.K., Gupta, R.D., Kumar, A. and Singh, B., 2011. Supplier issues for lean implementation. International Journal of Engineering Science and Technology, 3(5), pp.3900-3905.

Sivarajah, U., Kamal, M.M., Irani, Z. and Weerakkody, V., 2017. Critical analysis of Big Data challenges and analytical methods. Journal of Business Research, 70, pp.263-286.

Souza, G.C., 2014. Supply chain analytics. Business Horizons, 57(5), pp.595-605.

Stojanovic, L., Dinic, M., Stojanovic, N., and Stojadinovic, A. 2016, December. Big-data-driven anomaly detection in industry (4.0): An approach and a case study. In 2016 IEEE International Conference on Big Data (Big Data) (pp. 1647-1652). IEEE.

Stojanovic, N. and Milenovic, D., 2018, December. Data-driven Digital Twin approach for process optimization: an industry use case. In 2018 IEEE International Conference on Big Data (Big Data) (pp. 4202-4211). IEEE.

Strozzi, F., C. Colicchia, A. Creazza, and C. Noè. 2017. "Literature Review on the 'Smart Factory' Concept using Bibliometric Tools." International Journal of Production Research 55 (22): 6572–6591.

Sturm, C., Fichtner, M., and Schönig, S. 2019. Full support for efficiently mining multi-perspective declarative constraints from process logs. Information, 10(1), 29.

Sundar, R., Balaji, A.N. and Kumar, R.S., 2014. A review on lean manufacturing implementation techniques. Procedia Engineering, 97, pp.1875-1885.

Susto, G. A., Schirru, A., Pampuri, S., McLoone, S., and Beghi, A. 2014. Machine learning for predictive maintenance: A multiple classifier approach. IEEE Transactions on Industrial Informatics, 11(3), 812-820.

Syreyshchikova, N.V., Pimenov, D.Y., Mikolajczyk, T. and Moldovan, L., 2020. Automation of Production Activities of an Industrial Enterprise based on the ERP System. Procedia Manufacturing, 46, pp.525-532.

Talia, D., Trunfio, P. and Marozzo, F., 2015. Data analysis in the cloud: models, techniques and applications. Elsevier.

Tan, K.H., Zhan, Y., Ji, G., Ye, F. and Chang, C., 2015. Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. International Journal of Production Economics, 165, pp.223-233.

Tang, J. and Yan, X., 2017. Neural network modeling relationship between inputs and state mapping plane obtained by FDA–t-SNE for visual industrial process monitoring. Applied Soft Computing, 60, pp.577-590.

Tezel, A., Koskela, L., Tzortzopoulos, P., Formoso, C.T. and Alves, T., 2015. Visual management in Brazilian construction companies: taxonomy and guidelines for implementation. Journal of management in engineering, 31(6), p.05015001.

Tjahjono, B., Ball, P., Vitanov, V. I., Scorzafave, C., Nogueira, J., Calleja, J., Minguet, M., Narasimha, L., Rivas, A., Srivastava, A., Srivastava, S., and Yadav, A. (2010). Six Sigma: a literature review. International Journal of Lean Six Sigma, 1 (3), pp. 216-233.

Toasa, R., Maximiano, M., Reis, C., and Guevara, D. (2018, June). Data visualization techniques for real-time information—A custom and dynamic dashboard for analyzing surveys' results. In 2018 13th Iberian Conference on Information Systems and Technologies (CISTI). pp. 1-7. IEEE.

Tortorella, G. L., and Fettermann, D. 2018. Implementation of Industry 4.0 and lean production in Brazilian manufacturing companies. International Journal of Production Research, 56(8), 2975-2987.

Tortorella, Guilherme Luz, Ricardo Giglio, and Desirée H Van Dun. 2019. Industry 4.0 Adoption as a Moderator of the Impact of Lean Production Practices on Operational

Performance Improvement. International Journal of Operations & Production Management. Emerald Publishing Limited.

Valamede, L.S. and Akkari, A.C.S., 2020. Lean 4.0: A New Holistic Approach for the Integration of Lean Manufacturing Tools and Digital Technologies.

Van der Aalst, W. 2011. Process Mining: Discovery, Conformance and Enhancement of Business Processes. Springer-Verlag, Berlin.

Van der Aalst, W. and Damiani, E., 2015. Processes meet big data: Connecting data science with process science. IEEE Transactions on Services Computing, 8(6), pp.810-819.

van der Aalst, W.M., 2018. Process discovery from event data: Relating models and logs through abstractions. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(3), p.e1244.

Vicario, Grazia, and Shirley Coleman. 2020. “A Review of Data Science in Business and Industry and a Future View.” Applied Stochastic Models in Business and Industry 36 (1). Wiley Online Library: 6–18.

Vitari, C. and Raguseo, E., 2020. Big data analytics business value and firm performance: linking with environmental context. International Journal of Production Research, 58(18), pp.5456-5476.

Walha, A., Ghozzi, F. and Gargouri, F., 2019. From user generated content to social data warehouse: processes, operations and data modelling. International Journal of Web Engineering and Technology, 14(3), pp.203-230.

Waller, M.A. and Fawcett, S.E., 2013. Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. Journal of Business Logistics, 34(2), pp.77-84.

Waltman, L., N. J. van Eck, and E. C. Noyons. 2010. “A Unified Approach to Mapping and Clustering of Bibliometric Networks.” Journal of Informetrics 4 (4): 629–635.

Wamba, S.F., Akter, S., Edwards, A., Chopin, G. and Gnanzou, D., 2015. How 'big data'can make big impact: Findings from a systematic review and a longitudinal case study. International Journal of Production Economics, 165, pp.234-246.

Wang, B., J. Y. Zhao, Z. G. Wan, J. H. Ma, H. Li, and J. Ma. 2016. "Lean Intelligent Production System and Value Stream Practice." Paper Presented at the 3rd International Conference on Economics and Management, ICEM 2016, Jiangsu

Wang, Z., and Zhao, H. 2016, June. Empirical study of using big data for business process improvement at private manufacturing firm in cloud computing. In 2016 IEEE 3rd international conference on cyber security and cloud computing (CSCloud) (pp. 129-135). IEEE.

Watson, G. H., and DeYong, C. F. (2010). Design for Six Sigma: caveat emptor. International Journal of Lean Six Sigma, 1 (1), pp. 66-84.

Womack, J.P., Jones, D.T. and Roos, D., The machine that changed th e world. 1990. New York: Rawson Associates.

Woo, J.S., Suslow, P., Thorsen, R., Ma, R., Bakhtary, S., Moayeri, M. and Nambiar, A., 2019. Development and implementation of real-time web-based dashboards in a multisite transfusion service. Journal of pathology informatics, 10.

Xu, L. D., Xu, E. L., and Li, L. 2018. Industry 4.0: state of the art and future trends. International Journal of Production Research, 56(8), 2941-2962.

Xu, S., Zhang, X., Feng, L. and Yang, W., 2020. Disruption risks in supply chain management: a literature review based on bibliometric analysis. International Journal of Production Research, 58(11), pp.3508-3526.

Yan, S., and Yan, X. 2019. Using labeled autoencoder to supervise neural network combined with k-Nearest neighbor for visual industrial process monitoring. Industrial & Engineering Chemistry Research, 58(23), 9952-9958.

Yang, Hanna, Minjeong Park, Minsu Cho, Minseok Song, and Seongjoo Kim. 2014. “A System Architecture for Manufacturing Process Analysis Based on Big Data and Process Mining Techniques.” In 2014 IEEE International Conference on Big Data (Big Data), 1024–1029. IEEE.

Yao, Le, and Zhiqiang Ge. 2018. “Big Data Quality Prediction in the Process Industry: A Distributed Parallel Modeling Framework.” Journal of Process Control 68. Elsevier: 1–13.

Zhong, R.Y., Xu, C., Chen, C. and Huang, G.Q., 2017. Big data analytics for physical internet-based intelligent manufacturing shop floors. International journal of production research, 55(9), pp.2610-2621.

Zhu, Jinlin, Zhiqiang Ge, Zhihuan Song, and Furong Gao. 2018. “Review and Big Data Perspectives on Robust Data Mining Approaches for Industrial Process Modeling with Outliers and Missing Data.” Annual Reviews in Control 46. Elsevier: 107–133.

Zhuang, C., Liu, J., and Xiong, H. 2018. Digital twin-based smart production management and control framework for the complex product assembly shop-floor. The International Journal of Advanced Manufacturing Technology, 96(1-4), 1149-1163.

Zwetsloot, I. M., Kuiper, A., Akkerhuis, T. S., and de Koning, H. 2018. Lean Six Sigma meets data science: Integrating two approaches based on three case studies. Quality Engineering, 30(3), 419-431.