G Model
JMSY-632; No. of Pages 13
ARTICLE IN PRESS
Journal of Manufacturing Systems xxx (2018) xxx–xxx
Contents lists available at ScienceDirect
Journal of Manufacturing Systems
journal homepage: www.elsevier.com/locate/jmansys

Data-driven smart manufacturing

Fei Tao [a,*], Qinglin Qi [a], Ang Liu [b], Andrew Kusiak [c]

[a] School of Automation Science and Electrical Engineering, Beihang University, Beijing, 100191, China
[b] School of Mechanical and Manufacturing Engineering, University of New South Wales, Sydney, 2053, Australia
[c] Department of Mechanical and Industrial Engineering, The University of Iowa, Iowa City, USA

ARTICLE INFO

Article history:
Received 3 October 2017
Received in revised form 7 January 2018
Accepted 8 January 2018
Available online xxx

Keywords:
Big data
Smart manufacturing
Manufacturing data
Data lifecycle

ABSTRACT

The advances in the internet technology, internet of things, cloud computing, big data, and artificial intelligence have profoundly impacted manufacturing. The volume of data collected in manufacturing is growing. Big data offers a tremendous opportunity in the transformation of today's manufacturing paradigm to smart manufacturing. Big data empowers companies to adopt data-driven strategies to become more competitive. In this paper, the role of big data in supporting smart manufacturing is discussed. A historical perspective to data lifecycle in manufacturing is overviewed. The big data perspective is supported by a conceptual framework proposed in the paper. Typical application scenarios of the proposed framework are outlined.

1. Introduction

Manufacturers are embracing the notion of a convergence between the cyber and physical world. Manufacturing strategies have been developed, such as Industry 4.0 in Germany, Industrial Internet in the US, and the Made in China 2025 initiative. These programs promote the application of modern information technologies (new-IT) in manufacturing, which drives the development of smart manufacturing [1]. Smart manufacturing aims to convert data acquired across the product lifecycle into manufacturing intelligence in order to yield positive impacts on all aspects of manufacturing [2]. In the modern manufacturing industry, data generated by manufacturing systems is experiencing explosive growth, which has reached more than 1000 EB annually [3]. The systematic computational analysis of manufacturing data will lead to more informed decisions, which will in turn enhance the effectiveness of smart manufacturing [4]. In other words, data-driven manufacturing can be regarded as a necessary condition for smart manufacturing. Therefore, data is becoming a key enabler for enhancing manufacturing competitiveness [5], and manufacturers are beginning to recognize the strategic importance of data.

The value of big data does not hinge solely on the sheer volume of data under consideration, but rather on the information and knowledge that lies hidden in it. The emergence of New IT as the Internet of Things (IoT), cloud computing, mobile Internet, and artificial intelligence (AI), can be strategically leveraged and effectively integrated in support of data-driven manufacturing. For example, a number of innovative IoT solutions [6,7] promote the deployment of sensors in manufacturing to collect real-time manufacturing data. Cloud computing [8,9] enables networked data storage, management, and off-site analysis. Analysis results can be easily accessed by users through various mobile devices [10]. Artificial Intelligence (AI) solutions enable "smart" factories to make timely decisions with minimal human involvement [11].

Efforts to explore the applicability of big data in manufacturing have been initiated. A number of studies examining big data in manufacturing, including industrial automation [12], have emerged in recent years. Big data as a driver of industrial competitiveness was investigated in [13]. Dubey et al. [14] illustrate the unique role of big data analytics in sustainable manufacturing. Zhang et al. [15] propose a big data analytics architecture for clean manufacturing and maintenance processes. Other researchers have explored the role of big data in equipment maintenance [16], fault detection [17], fault prediction [18], and cost estimation [19], etc. In light of the inborn intelligence of big data, manufacturing systems must be made more "smart" to achieve the all-round monitoring, simulation, and optimization of production activities.

The rest of this paper is organized as follows. The evolvement history of manufacturing data is reviewed in Section 2. The lifecycle of manufacturing data is discussed in Section 3. The revolutionizing paradigm of big data driven smart manufacturing is presented

* Corresponding author at: School of Automation Science and Electrical Engineering, Beihang University, Beijing, 100191, China.
E-mail addresses: ftao@buaa.edu.cn (F. Tao), andrew-kusiak@uiowa.edu (A. Kusiak).

https://doi.org/10.1016/j.jmsy.2018.01.006

in Section 4, followed by an illustrative case study showcased in Section 5. Finally, conclusions are drawn in Section 6.

2. Historical perspectives on manufacturing data

As shown in Fig. 1, for a long time, information was documented on paper while manufacturing was realized by handicraft, therefore, the integration between information technology and manufacturing technology was neither beneficial nor feasible. Since the advent of ENIAC (i.e., the first electronic computer) in 1940s, the rapid development of information technology (IT) has been driving manufacturing toward informatization. The first numerical controlled (NC) milling machine was developed in the 1950s, which announced that manufacturing entered the NC era. Since the 1960s, the development of integrated circuits has paved the way for the advancement of computer hardware and software. Since the 1980s, TCP/IP, local area network (LAN), World Wide Web (WWW), and search engine emerged one after another to meet the increasing needs for data storage, indexing, processing, and exchange. All of these information technologies were widely applied in manufacturing. As a result, many advanced manufacturing technologies were put forward, such as computer integrated manufacturing (CIM), computer aided design (CAD), manufacturing execution system (MES), computer aided manufacturing (CAM), enterprise resource planning (ERP), and networked manufacturing (NM), etc. Recently, the rise of New IT (e.g., Internet of Things, cloud computing, big data analytics, and artificial intelligence) continues to revolutionize the manufacturing paradigm, leading to a series of new manufacturing concepts, for instance, manufacturing grid, cyber-physical manufacturing system, cloud manufacturing, etc. Due to the deep fusion between IT and manufacturing, the degree of manufacturing smartness is progressively elevated. As a result, the manufacturing data also becomes increasingly richer. The evolution of manufacturing data in four stages is discussed (see Fig. 1).

2.1. Data in the handicraft age

Prior to the First Industrial Revolution, the human society had been in the manual manufacturing stage for a long time. Artefacts were predominantly designed and manufactured by artisans [20]. As the most basic form of manufacturing, handicraft activities were of low complexity. As a result, the data generated in the production process was limited as it existed mostly in the form of human experience. In addition, experience was mostly transmitted verbally from one generation to the next, primarily based on apprenticeships. The key information and data could be easily lost, making production and quality control impossible to achieve. Due to the extremely low quantity and quality, the manufacturing data generated in the handicraft age was neither emphasized nor fully exploited. However, since handcrafting involves a high levels of human creativity, even today, it is used to manufacture luxury products (e.g., jewelry, watch, leather bag).

2.2. Data in the machine age

Generally speaking, the machine age consisted of two phases. As a result of the first industrial revolution, machines were employed as production tools in the early factories, leading to a significant increase in the scale of manufacturing. During this period, the relationship between humans and machines in production was highly complementary (i.e., early machines could only be operated by skilled operators to deliver their functions). Therefore, manufacturers began to emphasize two particular kinds of manufacturing data: worker-related data and machine-related data. Worker-related data (e.g., attendance, productivity, and performance) was used to facilitate decisions about issues such as salary structure, performance benchmarking, and work schedules. Machine-related data was used to support decisions concerning machine maintenance, repair, and replacement. Compared to the handicraft age, nevertheless, the First Industry Revolution introduced no significant changes to the way data was collected, stored, analyzed, transferred, and managed. As a matter of fact, workers still handled data manually based on empirical experience.

As a result of the Second Industrial Revolution (or the Technological Revolution), machine tools and interchangeable parts were widely incorporated into the "new" manufacturing process (e.g., the Bessemer process) in modern factories, leading to significant increases in manufacturing efficiency, and the manufacturing paradigm shifted to the mass production model [21]. The Second Industrial Revolution triggered some notable changes in the way data was processed. In particular, because of the division of work between managers and workers, manufacturing data was increasingly handled by educated managers. Moreover, managers began to employ more systematic methods to document and analyze manufacturing data. The raw data was extensively recorded in written documents (e.g., instructions, logbooks, notes, and charts) rather than stored in human memory. Scientific methods were used to determine the dependency relationships between different datasets. During this period, manufacturers began to exploit manufacturing data for cost reduction, quality control, and inventory management. In particular, statistical models were introduced to analyze a variety of quality-related manufacturing data, such as production planning, throughput yield, product quality, failure rate, raw material consumption, and scrap rate.

In summary, in the machine age, although a larger quantity of manufacturing data was analyzed through scientific methods, data was still handled manually by human operators (i.e., managers), as opposed to computers. Therefore, the utilization rate of manufacturing data remained relatively low.

2.3. Data in the information age

In the information age (or the digital age), information technologies were widely applied in manufacturing processes. As a consequence, the quantity of manufacturing data that companies were able to collect grew exponentially. A number of factors contributed to this growth in data. First, information systems (e.g., CRM, MES, ERP, SCM, PDM, etc.) were widely employed by manufacturers to facilitate production management. Second, computer systems (such as CAD, CAE, CAM, and FEA) were widely used to aid the creation, simulation, modification, and optimization of new products as well as manufacturing processes. Third, industrial robots and automatic machinery were commonly used in modern factories. More and more, electronic devices and digital computers were employed to automatically control production equipment. The evolvements in information technologies paved the way for manufacturers to achieve meeting customer needs better, quicker, and cheaper [22].

In the information age, data was stored in computer systems and managed by information systems. For example, customer data (e.g., home address, phone number, demographics), sales data (e.g., type, quantity, price, and shipping date of finished products), supply chain data (e.g., type, quantity, price and supplier of raw materials), financial data (e.g., assets, real property, tangible property, utility, intangible property, etc.), production planning data, bill of materials, inventory data (e.g., type, quantity, location of material and finished products in the warehouse), and maintenance data are all managed by CRM, MES, ERP, SCM, PLM, etc. Therefore, it could be easily exchanged among different departments or organizations. The efficiency of data analysis was significantly enhanced due to the use of computational models, although analysis results

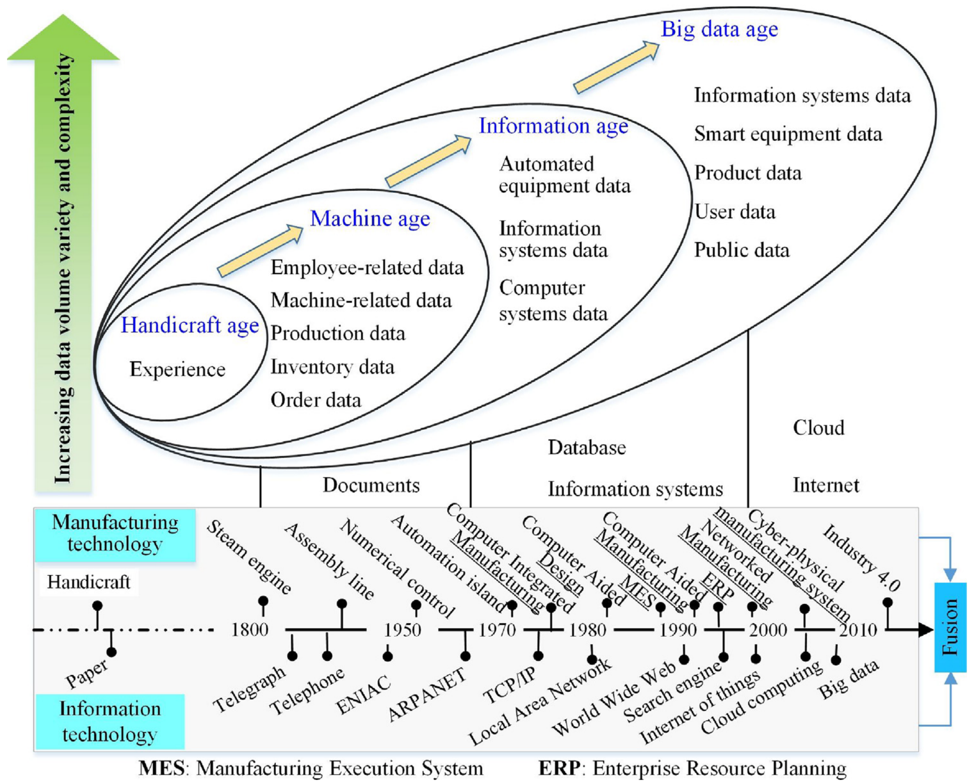

Fig. 1. Evolution of data in manufacturing.

still needed to be interpreted by human operators in order to make decisions. During this period, manufacturers began to leverage data to promote some advanced manufacturing models, such as mass customization, sustainable manufacturing, flexible manufacturing, intelligent manufacturing, and cloud manufacturing. Nevertheless, information silos (information systems that cannot communicate with other systems) were still common. There were no effective ways to analyze unstructured, scattered, repetitive, and isolated data. As a result, it was still difficult, especially for small- and medium-sized manufacturing enterprises, to benefit from the value of data.

2.4. Data in the big data age

Along with the rise of IoT technologies, cloud computing, big data analytics, AI, and other technological advances, came the age of big data [23]. In manufacturing, big data refers to large amounts of multi-source, heterogeneous data generated throughout the product lifecycle [24], which is characterized by 5 Vs [25], i.e., high volume (i.e., huge quantities of data), variety (i.e., the data itself comes in different forms and is generated by diverse sources), velocity (i.e., the data is generated and renewed at very high speed), veracity (i.e., the data is associated with a level of bias, inconsistency, incompleteness, ambiguities, latency, noises, and approximation), and value (i.e., huge value hidden in the data). Generally speaking, big data generated by manufacturing processes can be classified according to the following categories:

a) Management data collected from manufacturing information systems (e.g., MES, ERP, CRM, SCM, and PDM). Information systems possess a variety of data that is related to product planning, order dispatch, material management, production planning, maintenance, inventory management, sales and marketing, distribution, customer service, and financial management.
b) Equipment data collected from smart factories by Industrial IoT technologies, which includes data related to real-time performance, operating conditions, and the maintenance history of production equipment.
c) User data collected from internet sources such as ecommerce platforms (e.g., Amazon, Walmart, and Taobao) and social networking platforms (e.g., Twitter, Facebook, LinkedIn, and YouTube). This type of data encompasses user demographics, user profiles, user preferences towards products/services, as well as user behavior (e.g., data about online browsing, searching, purchasing, and reviewing history).
d) Product data collected from smart products and product-service systems by IoT technologies, including product performance, context of use (e.g., time, location, and weather), environmental data (e.g., temperature, humidity, and air quality) and user biological data.
e) Public data collected from governments through open databases. Such datasets deal with data related to intellectual property, civic infrastructure, scientific development, environmental protection, and health-care. For manufacturers, public data can be used to guarantee that manufacturing processes and manufactured products strictly comply with government regulations and industry standards.

In the big data age, empowered by the New ITs, manufacturer's ability to collect, store and process data is significantly enhanced. Recently, there emerged a number of cost-effective and flexible data collection, storage, and processing solutions such as the Internet of Things and cloud computing. As a result, manufacturing enterprises of different scales, even including SMEs, can benefit from the value of data. In manufacturing, effective analysis of big data enables manufacturers to deepen their understanding of customers, competitors, products, equipment, processes, services, employees, suppliers, and regulators. Therefore, big data can help

Please cite this article in press as: Tao F, et al. Data-driven smart manufacturing. J Manuf Syst (2018), https://doi.org/10.1016/j.jmsy.2018.01.006

manufacturers to make more rational, responsive, and informed decisions, and enhance their competitiveness in the global market.

The comparison of manufacturing data in different ages is shown in Table 1.

3. Lifecycle of manufacturing data

Data is a key enabler for smart manufacturing. However, data is not useful unless it is "translated" into concrete information content and context that can be directly understood by users [26]. Generally, before getting the concrete information from data, the data needs to pass through multiple steps. The complete journey of data collection, transmission, storage, pre-processing, filtering, analysis, mining, visualization, and application can be referred to the "data lifecycle" [27]. Manufacturing data is exploited at various points in the data *lifecycle*. As illustrated in Fig. 2, a typical manufacturing data lifecycle consists of data collection, transmission, storage, processing, visualization, and application.

3.1. Data sources

The volume of data collected across the entire manufacturing value-chain and product lifecycle is increasing at an unprecedented rate. As discussed in Section 2.4, the manufacturing data comes from equipment, products, human operators, information systems, and networks.

3.2. Data collection

Data from different sources is collected in a variety of ways. Above all, it is collected by means of the IoT, whereby equipment and product data can be instantly collected through smart sensors, RFID (radio frequency identification), and other sensing devices, making it possible to monitor equipment and product health in real time [28,29]. For instance, built-in sensors make it possible to continuously measure, monitor, and report the ongoing operational status of manufacturing equipment and products, such as temperature, pressure, and vibration. RFID enables the automatic identification, tracking, and management of a large number of workpieces, as well as the materials necessary for production. Moreover, the emerging mobile Internet paves the way for user data collection through smart terminals (e.g., devices like PCs, phones, laptops, and tablets). Through SDKs (software development kits) or APIs (application programming interfaces), for example, basic user data can be collected, including the number of users, user profiles, location, and time. In addition, web crawling [30] is a widely used data acquisition technique for collecting public data based on certain conditions predefined by engineers and AI. Web crawling refers to the technology of deploying "crawlers" (i.e., computer programs) to browse public web pages and collect desirable information. The web crawling technology enables manufacturers to acquire public data in an automatic and efficient manner. Last but not least, management data from manufacturing information systems can be acquired and used at any time through database technologies.

3.3. Data storage

The large volume of collected data from manufacturing processes must be securely stored and effectively integrated. Generally speaking, the various types of manufacturing data can be classified into structured (e.g. digit, symbols, tables, etc.), semi-structured (e.g., trees, graphs, XML documents, etc.), and unstructured data (e.g., logs, audios, videos, images, etc.) [31]. Traditionally, manufacturing enterprises focused heavily on structured data storage, since it was difficult to directly manage unstructured data within enterprise databases. Object-based storage architecture enables collections of data to be stored and managed as objects; this provides a more flexible solution for integrating semi-structured and unstructured data [32]. Also, through cloud computing [33], data storage can be achieved in a highly cost effective, energy efficient, and flexible fashion. In addition, by virtue of cloud services, the distribution and heterogeneity of data are shielded, enabling a highly scalable and shareable mode of data storage.

3.4. Data processing

Data processing refers to a series of operations conducted to discover knowledge from a large volume of data. Data must be converted to information and knowledge for manufacturers to make informed and rational decisions. Above all, data must be carefully preprocessed to remove redundant, misleading, duplicate, and inconsistent information. Specifically, data cleaning involves the following activities: missing value, format, duplicate, and garbage data cleaning. Data reduction is the process of transforming the massive volume of data into ordered, meaningful, and simplified forms by means of feature or case selection [34]. After data reduction had been completed, the cleaned and simplified data is exploited through data analysis and mining to generate new information. The effectiveness of data analysis can be significantly enhanced through a variety of techniques, including machine learning, large-scale computing, and the use of forecasting models. Some advanced data mining methods include clustering, classification, association rules, regression, prediction, and deviation analysis [27]. Through the above data processing efforts, understandable knowledge can be derived from a large quantity of dynamic and ambiguous raw data [35].

3.5. Data visualization

Visualization is intended to clearly convey and communicate information through graphical means, enabling end users to comprehend data in a much more explicit fashion [10]. The most commonly used visualization techniques include statement, chart, diagrams, graphs, and virtual reality [36]. Real-time data can be visualized online via users' smart terminals. Through visualization, the results of data processing are made more accessible, straightforward, and user-friendly.

3.6. Data transmission

Data is continuously flowing among different information systems, cyber-physical systems, and human operators. Data transmission, therefore, plays a critical role in maintaining communications and interactions among distributed manufacturing systems and resources. The recent advances in IoT, Internet, and communication networks substantially consolidated the technological foundation of real-time, reliable, and secure transmission of different types of data. As a result, distributed manufacturing resources can be effectively integrated almost anytime and anywhere.

3.7. Data applications

Data has entered almost all aspects of daily production and operation in manufacturing enterprises [37]. First, during the design phase, through data analytics, new insights are revealed about customers, competitors, and markets. Based on the understanding developed through data analytics, designers can accurately and rapidly translate customer voices to product features and quality requirements [38]. As a result, manufacturers will become "closer" to customers, and agiler in terms of coping with a dynamic, changing market. Second, during production, the manufacturing process

Table 1
Comparison of manufacturing data in different manufacturing ages.

	Data Source	Data Collection	Data Storage	Data Analysis	Date Transfer	Data Management
Handicraft Age	Human experience	Manual collection	Human memory	Arbitrary	Verbal communication	N/A
Machine Age	Human and machines	Manual collection	Written documents	Systematic	Written documents	Human operators
Information Age	Human, machines, information and computer systems	Semi-automated collection	Databases	Conventional algorithms	Digital files	Information systems
Big Data Age	Machines, product, user, information systems, public data	Automated collection	Cloud services	Big data algorithms	Digital files	Cloud and AI

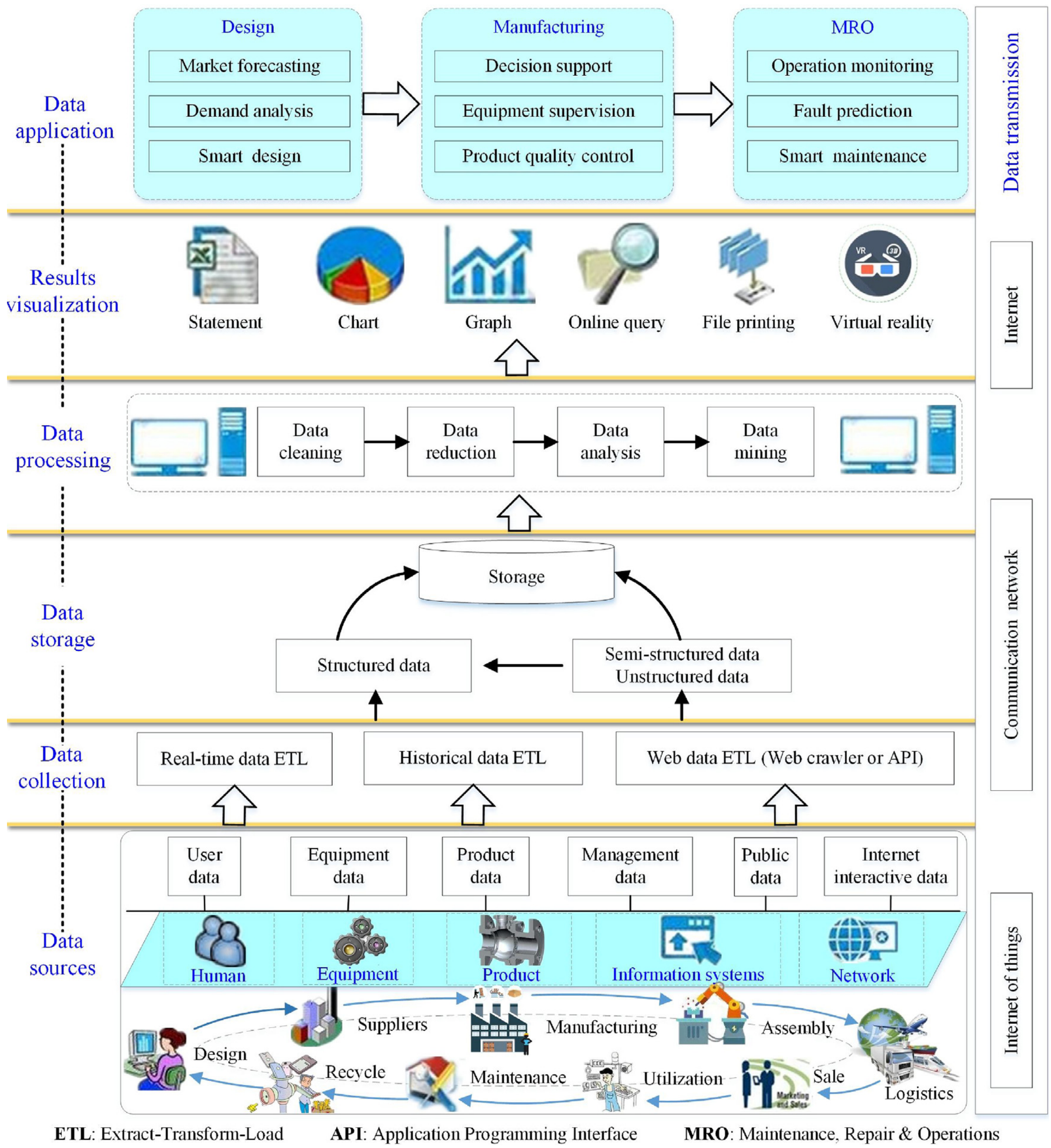

Fig. 2. Manufacturing data lifecycle.

and equipment are monitored and tracked in real time. In this way, the manufactures can keep abreast of changes. Data analytics can lead to informed decisions concerning whether, when, and how to adjust manufacturing processes and equipment. Additionally, data can facilitate the control and improvement of product quality. Data analytics can provide early warnings of quality defects and rapid diagnosis of root causes, both of which can be rapidly determined. Accordingly, manufacturing systems can be adjusted in a timely manner to control product quality. Lastly, with respect to product utilization and MRO, potential product malfunctions can be identified at an early stage [39], which makes precautionary actions possible, such as preventive maintenance, fault prediction, and automatic upgrade. For instance, through the development of prediction models, analysis of historical data can be used to predict the fault occurrence [40].

4. Data-driven smart manufacturing

4.1. The connotations of data-driven smart manufacturing

Manufacturing enterprises utilize big data analytics to exploit the data from manufacturing to refine manufacturing process, improving the flexibility and smart level of manufacturing. By taking full advantage of manufacturing data, manufacturing is shifted from primary processes to smart processes, thus improving the production efficiency and the performance of a product.

4.1.1. Data-driven smart manufacturing framework

The manufacturing data is collected, stored, processed, and analyzed by means of big data technologies. As a result, the degree of manufacturing intelligence can be significantly elevated.

As shown in Fig. 3, the data-driven smart manufacturing framework consists of four modules, namely, the manufacturing module, the data driver module, the real-time monitor module, and the problem processing module.

a) Manufacturing module: this module accommodates different kinds of manufacturing activities. It consists of a variety of information systems and manufacturing resources, which can be summarized as man-machine-material-environment. The inputs to this module are raw materials, whereas the outputs are finished products. During the input-output transformation process, various data is collected from human operators, production equipment, information systems, and industrial networks.
b) Data driver module: this module provides the driving force for smart manufacturing throughout the different stages of the manufacturing data lifecycle. As inputs, the data from the manufacturing module is transmitted to cloud-based data centers to be further analyzed. Afterwards, explicit information and actionable recommendations exploited from different kinds of raw data are used to direct the actions (e.g., product design, production planning, and manufacturing execution) in the manufacturing module. The real-time monitoring module and problem-processing module are also both powered by the data driver module.
c) Real-time monitoring module: this module plays a role in monitoring the manufacturing process in real time in order to ensure product quality. Driven by the data driver module, this module is enabled to analyze the real-time running status of manufacturing facilities. As a result, manufacturers can keep abreast of changes in the manufacturing process, so as to develop the optimal operational control strategies. For example, when a machine is idling, material is distributed and a trajectory is tracked. The manufacturing process can be adjusted in correspondence to specific product quality defects. As a result, the real-time monitoring module can make the manufacturing facilities run more efficiently.
d) Problem processing module: this module functions to identify and predict emerging problems (e.g., equipment faults or quality defects), diagnose root causes, recommend possible solutions, estimate solution effectiveness, and evaluate potential impacts on other manufacturing activities. Based on real-time information and analysis of historical and ongoing data provided by the data driver module, either human operators or artificial intelligence applications can make informed decisions, not only to address current problems, but also to prevent similar problems from happening in the future. The proactive maintenance enabled by this module will enhance smooth functioning of manufacturing processes.

The structured process of data collection, integration, storage, analysis, visualization and application is generally applicable for a variety of different industries. In that regard, the proposed data-driven smart manufacturing framework is intended to be universally valuable. With respect to the distinction between SMEs and big companies, depending on the resource availability, they can choose different strategies to achieve the data-driven smart manufacturing in different scales. For example, unlike those bigger companies that can afford to build an exclusive cloud infrastructure for data storage and analysis, SMEs can employ on-demand cloud computing services that are provided by third parties such as Amazon and Alibaba. Regardless where and how data is processed, the key value propositions of data-driven manufacturing are essentially the same for both SMEs and big companies. Manufacturing data helps decision makers understand changes in the shortest possible time, make accurate judgments regarding them, and develop rapid response measures to troubleshoot issues. As a consequence, production plans, manufacturing activities, and resources can be closely coordinated to promote smart manufacturing.

4.1.2. Characteristics of data-driven smart manufacturing

The data-driven smart manufacturing shares the following five characteristics (see Fig. 4):

(1) It enables customer-centric product development by exploiting user data for customized product design. For instance, user demographics, demands, preferences, and behaviors can be precisely quantified using big data analytics, so that more personalized products and services can be designed.
(2) It enables self-organization by exploiting manufacturing resources and task data for smart production planning. For instance, production plans can be created based on both internal and external data from different manufacturing sites. The appropriate manufacturing resources are chosen to form the optimal configuration, which meets all of the demands of the manufacturing task to implement production plans.
(3) It enables self-execution by exploiting a variety of data from the manufacturing process for precise control. For instance, appropriate raw material and parts can be sent to any manufacturing site that requires them at any time, and manufacturing equipment can automatically machine raw material or assemble parts where necessary.
(4) It enables self-regulation by exploiting real-time status data for manufacturing process monitoring. For instance, a manufacturing system can automatically respond to unexpected events (e.g., a shortage of manufacturing resources or a change in manufacturing tasks), by making its behaviors controllable, not only by human operators but also through AI systems.
(5) It enables self-learning and self-adaption by exploiting historical and real-time data for proactive maintenance and quality control. For instance, machine faults and quality defects can be predicted and prevented before they occur so that manufacturing systems can proactively adapt to cope with potential issues.

In summary, data-driven smart manufacturing provides a full range of services to manufacturing enterprises. One of the most important benefits is the ability to enable significant increases in manufacturing efficiency and remarkable improvements in product performance. Taking into account the characteristics outlined above and the manufacturing data lifecycle, the paradigm of data-

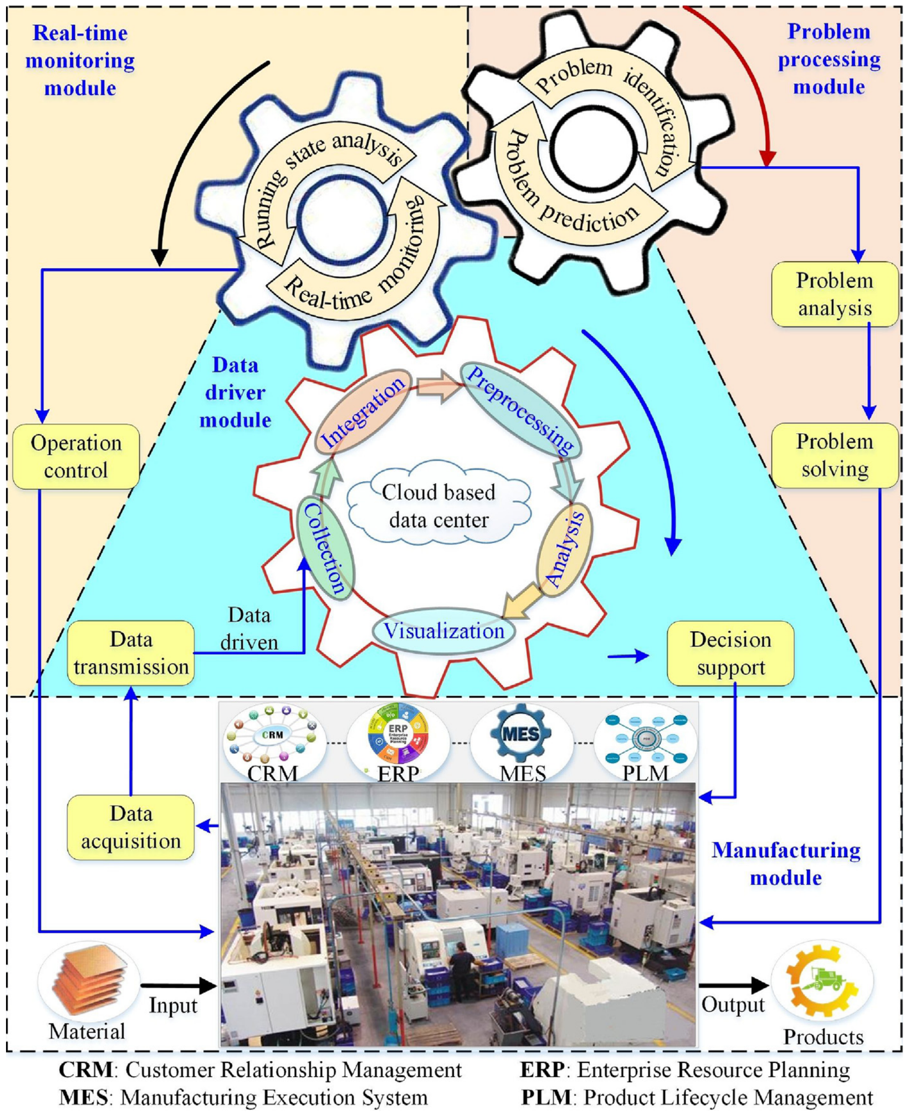

Fig. 3. The framework of data-driven smart manufacturing.

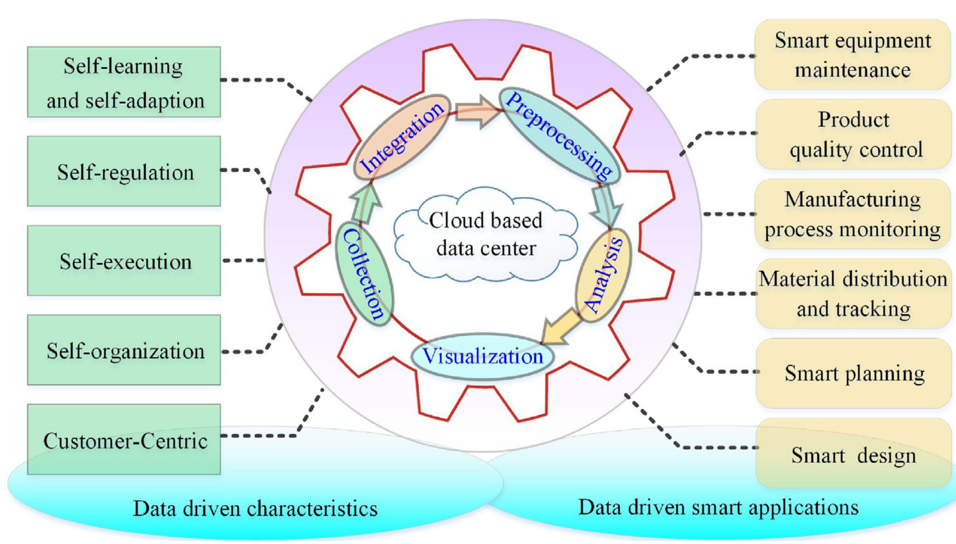

Fig. 4. Characteristics and applications of data-driven smart manufacturing.

driven smart manufacturing can be best exemplified through specific applications.

4.2. Data-driven-smart manufacturing application

Manufacturing converts raw material inputs into finished product outputs and value-added services through the coordination of relevant manufacturing facilities, resources, and activities. Some of the most promising applications that can be implemented during the manufacturing process include applications to enable smart design, smart planning, materials distribution and tracking, manufacturing process monitoring, quality control, and smart equipment maintenance (see Fig. 5).

4.2.1. Smart design

The importance of design cannot be overstated, since it determines most of a product's manufacturing costs. In the big data era, product design is shifting towards data-driven design [41]. Product design begins by researching and understanding customer demands, behaviors, and preferences. This type of data can be collected from both Internet and IoT sources. In the case of Internet data, customers are becoming increasingly good at sharing their first-hand experiences of using a product on the Internet, through portals such as social networking sites, ecommerce platforms, and product/service review sites [42]. In the case of IoT sources, rich user data (e.g., biological data, behavior data, and user-product interactions) can now be gathered from a growing number of increasingly popular smart products (e.g., smartphones and wearable devices) that are connected to IoT infrastructure. The holistic consideration of harnessing user-related big data improves the capacity for manufacturers to translate customer voices into product features and quality requirements. In addition, it enables designers to streamline design processes, promote product innovations, and develop more customized products for end users [43] [44]. Moreover, compared to traditional methods (e.g., interviews, surveys, etc.), in virtue of cloud-based high performance computing, big data analytics enables users to not only accelerate computationally expensive tasks (e.g., market preferences and customer demands analysis, etc.), but also reduce costs [45].

4.2.2. Smart planning and process optimization

Even before manufacturing of a product begins, production planning is necessary to determine the production capacity of a manufacturing facility, as well as the availability of resources and materials. Big data analytics can make production planning and shop floor scheduling more intelligent [46]. First of all, a variety of data, such as customer orders, manufacturing resource status, production capacities, supply chain data, sales data, and inventory data is analyzed using big data analytics methods. Based on the information gathered from these approaches, hypernetwork based manufacturing resource supply-and-demand matching and scheduling [47] can be carried out to rapidly locate available resources. Next, production plans are developed using intelligent optimization algorithms to determine the optimal configuration of manufacturing resources and the execution procedures for the task [48,49]. In addition, process optimization is also an important consideration before manufacturing begins. Big data analytics aid in assessing and optimizing technological processes. By analyzing various types of process data, including historical data and data on the patterns and relationships inherent to particular processing steps, the correlation between different technological parameters and the effect of these parameters on yield and quality can be determined. Adjusting technological processes in relation to these parameters can result in improved productivity and product quality, as well as reduced costs.

4.2.3. Material distribution and tracking

Material distribution is determined through production planning and actual production progress, as well as various on-site urgency requirements. In the ideal scenario, the right material should be delivered to right equipment at the right time, so that it can be processed through the right operations. To support this ideal, a variety of material-related data, including inventory data, logistics data, and progression data can be managed [50]. Material data is analyzed in association with multi-source data related to material flow (e.g., data from human operators, machines, vehicles, etc.). In performing these analyses, material distribution can be determined in terms of material kind, quantity, delivery time and method in order to support optimal manufacturing logistics. For example, material can be dispatched on time, according to the actual production pace and conditions, to ensure smooth production (i.e., avoiding unnecessary production delays, interruptions, or production stoppages). Moreover, traceability of materials [51] is necessary to ensure that certain types of materials strictly comply with their corresponding quality criteria norms and standards. By deploying identification tags, material conditions (e.g., location, status, and quality) can be tracked in real time throughout the entire production process. For example, RFID-enabled positioning system in AGV enables the efficient delivery of material within the manufacturing sites [52]. Based on big data analytics, operational data conducive to product quality control and product defect traceability can be generated during production.

4.2.4. Manufacturing process monitoring

The manufacturing process consists of multiple manufacturing factors. These factors (e.g., manufacturing equipment, material, environment, and technological parameters) can affect the manufacturing process and influence changes in product quality. In addition, they can also interact with each other. Therefore, it is particularly important to monitor different steps of the manufacturing process in real time. However, it is often difficult to systematically trace which factors affect manufacturing processes. Fortunately, big data provides effective technical support for monitoring manufacturing processes. Assisted by the predictive capacity of big data analytics, the most suitable design range for each manufacturing factor can be prescribed. Once a factor falls outside its acceptable range, the problem will be flagged, and alerts and recommendations will be sent to operators to make timely adjustments; this can ensure greater uniformity in the manufacturing process. Taking the production abnormities in shop-floor for example, the abnormities (e.g., tardiness of order) are often caused by anomalous events, such as equipment failure, lack of material, and operation deviation, etc. Before the occurrence of production abnormities, the anomalous events often reveal certain patterns that can be captured by a variety of data (e.g., material consumption data, energy consumption data, rotation rate, vibration, torque, etc.) in time series. Since such data is mostly time-dependent, it cannot be effectively processed by means of static models [53]. Furthermore, the big data cannot be processed by traditional data analysis methods, which are computationally intractable [53]. By synthetizing the factors of time and causality [54], an early-warning model of production abnormities in shop-floor can be established based on relevant big data algorithms, for instance, decision tree (e.g., ID3 and C4.5) and neural network [55,56]. By mining the feature patterns and the trend of abnormal events in time series, it is possible to predict, in advance, whether and when production abnormities will occur. With higher flexibility, accuracy and less computing time, big data analytics can deal with multi-source data and massive data. Taking balanced use into consideration, manufacturing processes can be dynamically adjusted based on big data analytics.

Please cite this article in press as: Tao F, et al. Data-driven smart manufacturing. J Manuf Syst (2018), https://doi.org/10.1016/j.jmsy.2018.01.006

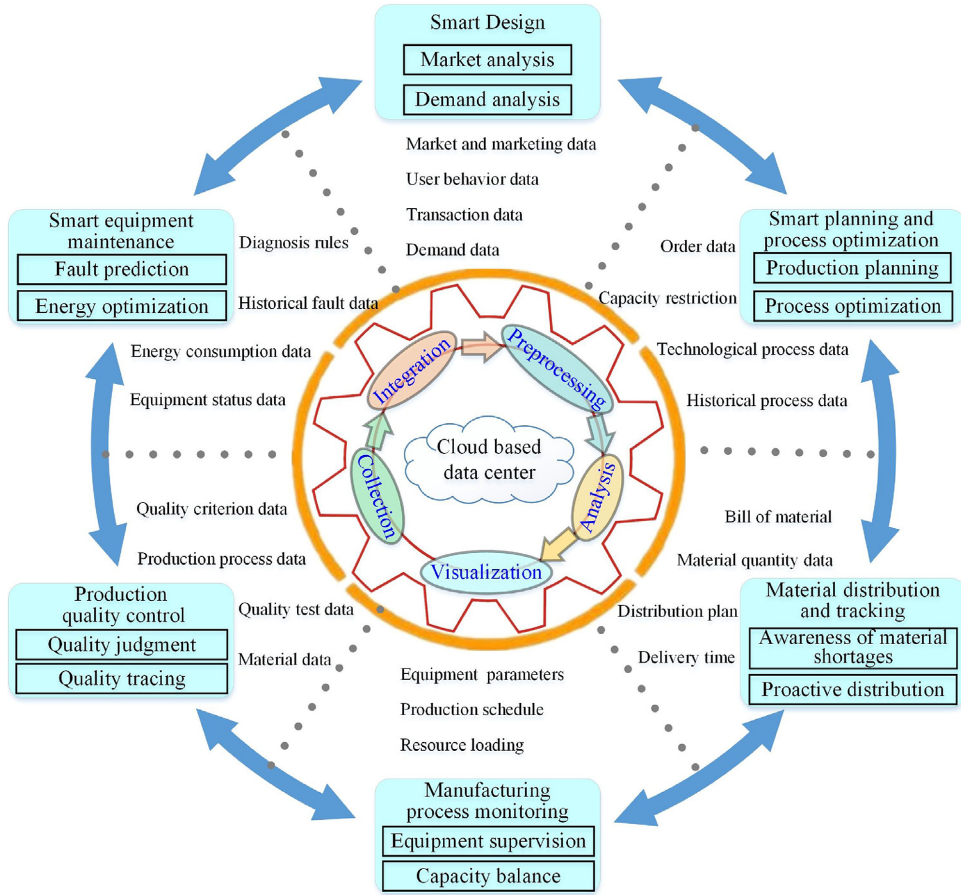

Fig. 5. Data-driven smart manufacturing application.

4.2.5. Product quality control

Various data-driven quality control techniques are being developed for smart manufacturing [57]. A variety of sensors, RFIDs and machine vision applications can be employed to collect product quality data, such as geometric parameters (e.g., thickness, length and surface roughness), location parameters (e.g., coordinate), tolerance parameters (e.g., concentricity), machining parameter (e.g., pressure, speed, temperature and machining time), etc. [24]. Big data analytics can serve the all-around quality monitoring, early warning of quality defects, and rapid diagnosis of root causes [58]. Based on historical data and process condition data gathered from machines and their operating environment, the binary classification of quality conditions can be used to predict whether and how certain conditions are related to quality defects [57]. Bayesian inference method can be used to analyze the data of process parameters and defective products to identify the most influential parameters and their appropriate range [59]. In addition, the root cause analysis together with the weighted association rule mining can be used to identify the root causes of product failures [60]. Thus, product quality defects can be detected, diagnosed, and addressed in a timely manner. In particular, less explicit causes of production issues, such as couplings between different equipment and inefficient procedures, can be illuminated by means of data integration and data mining. As a result, not only can low quality or failed products be automatically identified and removed, but factors that result in quality defects can also be eliminated or controlled. In addition, in conjunction with machine learning, big data analytics will eventually equip manufacturing enterprises with a particular kind of case-based reasoning capacity. Lessons learned from one quality control case can be transferred to another to prevent the recurrence of similar problems in the future. As a result, quality management can be embedded into every step of the manufacturing process, from raw materials to finished product.

4.2.6. Smart equipment maintenance

Data analytics can accurately predict and diagnose equipment faults and component lifetime [61,62]; such information can be used to enable informed maintenance decisions. In combination with the equipment status data from smart sensors - as well as domain knowledge, previous experience, and historical records concerning equipment maintenance - big data analytics can predict the tendency for equipment capacity to deteriorate, the lifespan of components, and the cause and extent of certain faults [46]. In addition, seasonal, periodic, combinational, and other patterns of equipment faults can also be discovered through big data analytics. With this information, precautionary actions can be performed to prevent faults. Because of the predictive capacity of big data analytics, the equipment maintenance paradigm is transformed from passive to proactive maintenance, thus prolonging equipment life and minimizing maintenance costs [63]. Energy consumption [64,65] is also an important reference for equipment faults or abnormalities. By establishing a multi-dimensional energy consumption analysis model, big data related to energy consumption can help to uncover energy fluctuations and abnormalities or peaks in real time. To ensure normal production, the corresponding production processes, equipment, and energy supplies can be dynamically adjusted to achieve optimization in real time.

5. Case study

In this section, a case study is presented to illustrate some practical aspects of the proposed framework. This case describes a silicon

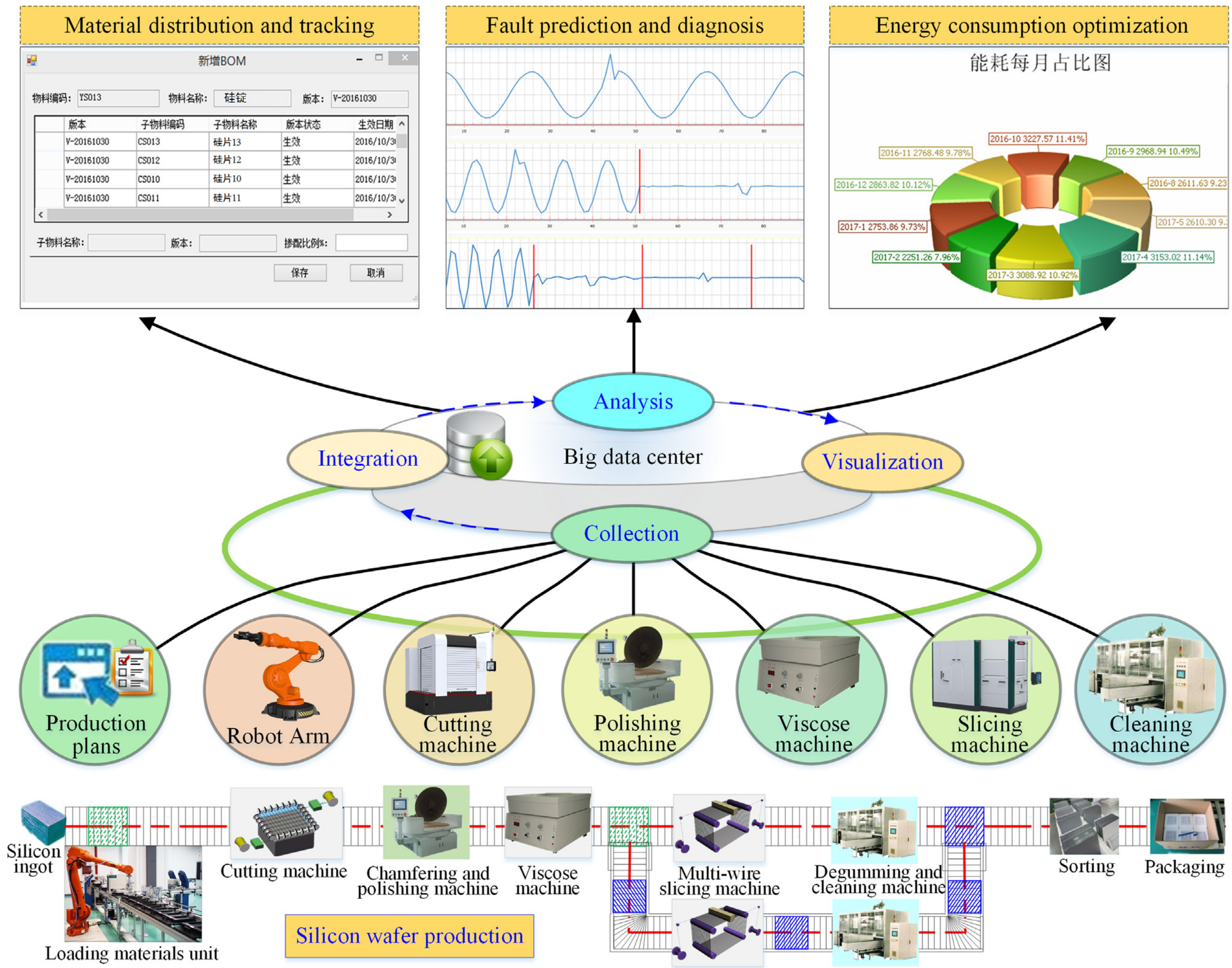

Fig. 6. Data-driven smart silicon wafer production.

wafer production line (illustrated in Fig. 6). Silicon wafers are one of the most important components of crystalline silicon photovoltaic cells, which play a critical role in improving solar energy products. As shown in Fig. 6., from the input of silicon ingots to the output of silicon wafers, the manufacturing process involves a series of production activities, including loading material, cutting, chamfering and polishing, viscos, slicing, degumming and cleaning, sorting, and packaging. Accordingly, the production line consists of multiple pieces of equipment associated with these processes.

As shown in Fig. 6, a variety of different types of multi-source and heterogeneous data generated in the production process are continuously accumulated. Data from the production process is integrated with information from orders and production plans. With the support of big data analytics, intelligent algorithms and predictive models analyze this data in order to optimize the manufacturing process. As a result, big data analytics enables intelligent material assignment, as well as tracking, predictive maintenance, and energy efficiency management.

For material distribution and tracking, RFID tags are embedded into materials, and external readers are deployed to collect material data. The material data is represented as 9-tuples, i.e., Material = {ID, time, location, batch, type, quantity, sender, receiver, item code}. Material data is collected every time when the material goes through each RFID reader. From the input of silicon ingots to the output of silicon wafers, a huge volume of data is generated. Material identification is achieved through the data fusion technology in accordance with the set single recognition confidence. If the actual result is lower than the set confidence, an alarm is sent to the operator through mobile terminals (e.g. smartphone and tablet computer) for manual processing. Furthermore, through analysis of the material data, the operator can monitor in real time where and how the material is being processed at any particular time point. Lastly, a dynamic material distribution scheme is developed to constantly capture the location, batch, type and quantity of to-be-delivered material. In addition, delivery instructions and routes are visualized for operators through mobile terminals. The whole material dataset, including when, where, and which production process the material was going through, can be retrieved and reviewed at any time.

Second, for the purpose of fault diagnosis and prediction, sensors are embedded in production equipment to detect a variety of data, including variables such as location, weight, temperature, humidity, vibration, and flow rate. Real-time data is used to determine which equipment requires service, repair, and even replacement. In this case study, for example, vibration data is used to diagnose running state of the multi-wire slicing machine by means of multiple vibration sensors. The vibration data can be leveraged to characterize the operational patterns of the multi-wire slicing machine. Firstly, the noisy and redundant data is removed through a denoising method based on wavelet transform module maxima. Next, the feature extraction method based on attribute reduction is used to extract feature parameters from the vibration data. Finally,

(a) (b)

Fig. 7. The software interfaces of energy consumption analysis.

based on the BP neural network, a smart failure diagnosis is performed. Specifically, under normal circumstances, the vibration signal should demonstrate a relatively stable pattern. When the equipment is worn or unexpected faults occur, the vibration signal would deviate from the normal pattern, which will automatically trigger a warning to be sent to the operator through mobile terminals. Through analysis of the vibration signal, equipment anomalies can be predicted and diagnosed.

Finally, with respect to energy efficiency management, energy consumed in the manufacturing process is measured by smart meters installed in each piece of production equipment. The data goes through a multi-dimensional analysis. Firstly, through the hierarchical cluster analysis, the energy consumption patterns for a fixed period of time are discovered in order to improve energy efficiency. One advantage of the hierarchical cluster analysis method is that it no longer requires the time-consuming modeling and computing. By analyzing the daily, weekly, monthly and yearly energy consumption, the change law of energy consumption over time can be formulated. The results of hierarchical cluster analysis can be visualized, as shown in Fig. 7(a). The predictive analytics of energy consumption is performed based on the autoregressive integrated moving average (ARIMA) algorithm. As shown in Fig. 7(b), the red curves represent the monthly, weekly and daily records of energy consumption, through which, manufacturers can clearly see the trends and characteristics of energy consumption in both short term and long term, and hence make energy plans accordingly. Moreover, by comparing real-time data with historical data, the overall patterns and/or trends of energy consumption changes can be evaluated; this information is useful for determining whether and when to conduct a comprehensive overhaul. In addition, unusual fluctuations in energy consumption can serve as additional indicators of abnormalities in production processes.

6. Conclusion and future work

The volumes of dynamically changing data generated throughout the lifecycle of products constitutes is growing. The data collected can be used to increase efficiency of manufacturing industry. This paper has provided contributions to smart manufacturing in three perspectives. (1) historical perspective: the evolution of manufacturing data was reflected in accordance with four manufacturing eras: the handicraft age, the machine age, the information age, and the big data age; (2) development perspective: the lifecycle of big manufacturing data was illustrated as a series of phases that includes data generation, collection, transmission, storage and integration, processing and analysis, visualization and application; (3) envisioning the future of data in manufacturing perspective: the role of data analytics in manufacturing was discussed, in particular with respect to promising applications in smart manufacturing.

There are multiple limitations that should be considered. Firstly, the current data collection technologies are not fully ready for smart data perception, especially when dealing with heterogeneous devices that are equipped with different communication interfaces and protocols [29]. Secondly, although the cloud-based data storage and analytics is proven to be a feasible technological solution, there remain some unresolved issues (e.g., network unavailability, overfull bandwidth, and unacceptable latency time, etc.,) that limit its applicability for the low-latency and real-time applications [66]. Thirdly, although it is commonly agreed that the integration between the physical and cyber worlds is a key feature of smart manufacturing, the vast majority of previous researches mainly focused on data collected from the physical world instead of data from virtual models [37]. That being said, this paper serves as a preliminary exploration of data-driven smart manufacturing and its potential applications. With respect to future work, there are some promising directions that can be pursued by interested researchers:

(1) The key technologies for data perception and collection from heterogeneous equipment, such as IoT gateways or industrial Internet hub) can be incorporated into the data-driven smart manufacturing framework. The devices that are compatible with the heterogeneous interfaces and communication protocols will be more conducive to data collection and data transmission.
(2) The new technologies for data storage and processing, such as fog computing and edge computing, can be incorporated into the proposed framework. Fog computing and edge computing can extend the manufacturer's data computing, storage, and networking capabilities from the cloud to the edge, which will significantly reduce bandwidth requirement, latency time, and service downtime [67].
(3) The digital twin technologies can be incorporated into the proposed framework. Digital twin enables manufacturers to manage the real-time, two-way, and coevolving mapping between a physical object and its digital representation, which paves the way for the deep cyber-physical integration. In combination with digital twin, the data-driven smart manufacturing will be made more responsive, adaptable, and predictive.

Acknowledgments

This work is financially supported in part by the Beijing Nova Program in China under Grant Z161100004916063, in part by National Natural Science Foundation of China (NSFC) under Grant

51522501, and in part by the National Key Research and Development Program of China under Grant 2016YFB1101700.

References

[1] Tao F, Qi Q. New IT driven service-oriented smart manufacturing: framework and characteristics. IEEE Trans Syst Man Cybern Syst 2017, http://dx.doi.org/10.1109/TSMC.2017.2723764.
[2] O'Donovan P, Leahy K, Bruton K, O'Sullivan DTJ. An industrial big data pipeline for data-driven analytics maintenance applications in large-scale smart manufacturing facilities. J Big Data 2015;2(25):1–26, http://dx.doi.org/10.1186/s40537-015-0034-z.
[3] Yin S, Kaynak O. Big data for modern industry: challenges and trends [point of view]. Proc IEEE 2015;103(2):143–6.
[4] Shao G, Shin SJ, Jain S. Data analytics using simulation for smart manufacturing. Proceedings of the 2014 Winter Simulation Conference 2014:2192–203.
[5] Kusiak A. Smart manufacturing must embrace big data. Nature 2017;544(7648):23–5.
[6] Mourtzis D, Vlachou E, Milas N. Industrial big data as a result of IoT adoption in manufacturing. Procedia CIRP 2016;55:290–5.
[7] Tao F, Zuo Y, Xu LD, Zhang L. IoT-based intelligent perception and access of manufacturing resource toward cloud manufacturing. IEEE Trans Ind Inf 2014;10(2):1547–57.
[8] Hashem IAT, Yaqoob I, Anuar NB, Mokhtar S, Gani A, Khan SU. The rise of big data on cloud computing: review and open research issues. Inf Syst 2015;47:98–115.
[9] Tao F, Cheng Y, Xu LD, Zhang L, Li BH. CCIoT-CMfg: cloud computing and internet of things-based cloud manufacturing service system. IEEE Trans Ind Inf 2014;10(2):1435–42.
[10] Lee MD, Butavicius MA, Reilly RE. Visualizations of binary data: a comparative evaluation. Int J Hum Comput Stud 2003;59(5):569–602.
[11] Wuest T, Weimer D, Irgens C, Thoben KD. Machine learning in manufacturing: advantages, challenges, and applications. Prod Manuf Res 2016;4(1):23–45.
[12] Obitko M, Jirkovský V, Bezdíček J. Big data challenges in industrial automation. Proceedings of International Conference on Industrial Applications of Holonic and Multi-Agent Systems 2013:305–16.
[13] Galletti A, Papadimitriou DC. How big data analytics are perceived as a driver for competitive advantage: A qualitative study on food retailers; 2013. p. 1–58. Master Thesis.
[14] Dubey R, Gunasekaran A, Childe SJ, Wamba SF, Papadopoulos T. The impact of big data on world-class sustainable manufacturing. Int J Adv Manuf Technol 2016;84(1–4):631–45.
[15] Zhang Y, Ren S, Liu Y, Si S. A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products. J Clean Prod 2017;142:626–41.
[16] Bahga A, Madisetti VK. Analyzing massive machine maintenance data in a computing cloud. IEEE Trans Parallel Distrib Syst 2012;23(10):1831–43.
[17] Lee H. Framework and development of fault detection classification using IoT device and cloud environment. J Manuf Syst 2017;43:257–70.
[18] Munirathinam S, Ramadoss B. Big data predictive analytics for proactive semiconductor equipment maintenance. Proceedings of 2014 IEEE International Conference on Big Data 2014:893–902.
[19] Chan SL, Lu Y, Wang Y. Data-driven cost estimation for additive manufacturing in cyber manufacturing. J Manuf Syst 2018;46:115–26.
[20] Schleich B, Anwer N, Mathieu L, Wartzack S. Shaping the digital twin for design and production engineering. CIRP Ann Manuf Technol 2017;66(2017):141–4.
[21] Hu SJ. Evolving paradigms of manufacturing: from mass production to mass customization and personalization. Procedia CIRP 2013;7:3–8.
[22] Balakrishnan A, Kumara SR, Sundaresan S. Manufacturing in the digital age: exploiting information technologies for product realization. Inf Syst Front 1999;1(1):25–50.
[23] Bughin J, Chui M, Manyika J. Clouds, big data, and smart assets: ten tech-enabled business trends to watch. McKinsey Q 2010;56(1):75–86.
[24] Li J, Tao F, Cheng Y, Zhao L. Big data in product lifecycle management. Int J Adv Manuf Technol 2015;81(1–4):667–84.
[25] Chen M, Mao S, Liu Y. Big data: a survey. Mobile Networks Appl 2014;19(2):171–209.
[26] Lee J, Lapira E, Bagheri B, Kao HA. Recent advances and trends in predictive manufacturing systems in big data environment. Manuf Lett 2013;1(1):38–41.
[27] Siddiqa A, Hashem IAT, Yaqoob I, Marjani M, Shamshirband S, Gani A, et al. A survey of big data management: taxonomy and state-of-the-art. J Network Comput Appl 2016;71:151–66.
[28] Zhang Y, Zhang G, Wang J, Sun S, Si S, Yang T. Real-time information capturing and integration framework of the internet of manufacturing things. Int J Comput Integr Manuf 2015;28(8):811–22.
[29] Tao F, Cheng J, Qi Q. IIHub: an industrial internet-of-Things hub towards smart manufacturing based on cyber-Physical system. IEEE Trans Ind Inf 2017, http://dx.doi.org/10.1109/TII.2017.2759178.
[30] Guerriero A, Ragni F, Martines C. A dynamic URL assignment method for parallel web crawler. Proceedings of 2010 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications 2010:119–23.
[31] Gandomi A, Haider M. Beyond the hype: big data concepts, methods, and analytics. Int J Inf Manage 2015;35(2):137–44.
[32] Nicolae B, Antoniu G, Bougé L, Moise D, Carpen-Amarie A. BlobSeer: next-generation data management for large scale infrastructures. J Parallel Distrib Comput 2011;71(2):169–84.
[33] Agrawal D, El Abbadi A, Antony S, Das S. Data management challenges in cloud computing infrastructures. Proceedings of 6th international workshop on databases in networked information systems 2010:1–10.
[34] Huang J, Li YF, Xie M. An empirical analysis of data preprocessing for machine learning-based software cost estimation. Inf Software Technol 2015;67:108–27.
[35] Begoli E, Horey J. Design principles for effective knowledge discovery from big data. Proceedings of 2012 Software Architecture and European Conference on Software Architecture 2012:215–8.
[36] Mittal S, Khan MA, Romero D, Wuest T. Smart manufacturing: characteristics, technologies and enabling factors Proceedings of the Institution of Mechanical Engineers. Part B: J Eng Manuf 2017, http://dx.doi.org/10.1177/0954405417736547.
[37] Tao F, Cheng J, Qi Q, Zhang M, Zhang H, Sui F. Digital twin-driven product design, manufacturing and service with big data. Int J Adv Manuf Technol 2017, http://dx.doi.org/10.1007/s00170-017-0233-1.
[38] Liu Y, Jin J, Ji P, Harding A, Fung RY. Identifying helpful online reviews: a product designer's perspective. Comput Aided Des 2013;45(2):180–94.
[39] Bennane A, Yacout S. LAD-CBM; new data processing tool for diagnosis and prognosis in condition-based maintenance. J Intell Manuf 2012;23(2):265–75.
[40] Kusiak A, Verma A. Analyzing bearing faults in wind turbines: a data-mining approach. Renew Energ 2012;48:110–6.
[41] Kusiak A, Salustri FA. Computational intelligence in product design engineering: review and trends. IEEE Trans Syst Man Cybern Part C (Appli Rev) 2007;37(5):766–78.
[42] Qi J, Zhang Z, Jeon S, Zhou Y. Mining customer requirements from online reviews: a product improvement perspective. Inf Manage 2016;53(8):951–63.
[43] Innovation Kusiak A. A data-driven approach. Int J Prod Econ 2009;122(1):440–8.
[44] Da Cunha C, Agard B, Kusiak A. Selection of modules for mass customisation. Int J Prod Res 2010;48(5):1439–54.
[45] Wu D, Liu X, Hebert S, Gentzsch W, Terpenny J. Democratizing digital design and manufacturing using high performance cloud computing: performance evaluation and benchmarking. J Manuf Syst 2017;43:316–26.
[46] Ji W, Wang L. Big data analytics based fault prediction for shop floor scheduling. J Manuf Syst 2017;43:187–94.
[47] Cheng Y, Tao F, Zhang L, Zhao D. Dynamic supply-demand matching for manufacturing resource services in service-oriented manufacturing systems: a hypernetwork-based solution framework. Proceedings of ASME International Manufacturing Science and Engineering Conference (MSEC2015) 2015;2015:8–12.
[48] Song Z, Kusiak A. Multiobjective optimization of temporal processes. IEEE Trans Syst Man Cybern Part B (Cybern) 2010;40(3):845–56.
[49] Tao F, Zhao D, Hu Y, Zhou Z. Resource service composition and its optimal-selection based on particle swarm optimization in manufacturing grid system. IEEE Trans Ind Inf 2008;4(4):315–27, 2008.
[50] Zhong RY, Lan S, Xu C, Dai Q, Huang GQ. Visualization of RFID-enabled shopfloor logistics big data in cloud manufacturing. Int J Adv Manuf Technol 2016;84(1–4):5–16.
[51] Zhong RY, Huang GQ, Lan S, Dai QY, Chen X, Zhang T. A big data approach for logistics trajectory discovery from RFID-enabled production data. Int J Prod Econ 2015;165:260–72.
[52] Lu S, Xu C, Zhong RY, Wang L. A RFID-enabled positioning system in automated guided vehicle for smart factories. J Manuf Syst 2017;44:179–90.
[53] Qin SJ. Survey on data-driven industrial process monitoring and diagnosis. Annu Rev Control 2012;36(2):220–34.
[54] Qin SJ. Process data analytics in the era of big data. AIChE J 2014;60(9):3092–100.
[55] Abonyi J, Roubos JA, Szeifert F. Data-driven generation of compact: accurate, and linguistically sound fuzzy classifiers based on a decision-tree initialization. Int J Approximate Reasoning 2003;32(1):1–21.
[56] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science 2006;313(5786):504–7.
[57] Kim A, Oh K, Jung JY, Kim B. Imbalanced classification of manufacturing quality conditions using cost-sensitive decision tree ensembles. Int J Comput Integr Manuf 2017, http://dx.doi.org/10.1080/0951192X.2017.1407447.
[58] Köksal G, Batmaz İ, Testik MC. A review of data mining applications for quality improvement in manufacturing industry. Expert Syst Appl 2011;38(10):13448–67.
[59] Sata A, Ravi B. Bayesian inference-based investment-casting defect analysis system for industrial application. Int J Adv Manuf Technol 2017;90(9–12):3301–15.
[60] He Y, Zhu C, He Z, Gu C, Cui J. Big data oriented root cause identification approach based on Axiomatic domain mapping and weighted association rule mining for product infant failure. Comput Ind Eng 2017;109:253–65.
[61] Verma A, Zhang Z, Kusiak A. Modeling and prediction of gearbox faults with data-mining algorithms. J Solar Energy Eng 2013;135(3), 031007-1-11.
[62] Si XS, Wang W, Hu CH, Zhou DH. Remaining useful life estimation–A review on the statistical data driven approaches. Eur J Oper Res 2011;213(1):1–14.

[63] Zhang Z, He X, Kusiak A. Data-driven minimization of pump operating and maintenance cost. Eng Appl Artif Intell 2015;40:37–46.
[64] Mourtzis D, Vlachou E, Milas N, Dimitrakopoulos G. Energy consumption estimation for machining processes based on real-time shop floor monitoring via wireless sensor networks. Procedia CIRP 2016;57:637–42.
[65] Zuo Y, Tao F, Nee AYC. An Internet of things and cloud-based approach for energy consumption evaluation and analysis for a product. Int J Comput Integr Manuf 2017, http://dx.doi.org/10.1080/0951192X.2017.1285429.
[66] Pizoń J, Jerzy L. Perspectives for fog computing in manufacturing. Appl Comput Sci 2016;12(3):37–46.
[67] Wu D, Liu S, Zhang L, Terpenny J, Gao RX, Kurfess T, et al. A fog computing-based framework for process monitoring and prognosis in cyber-manufacturing. J Manuf Syst 2017;43:25–34.