Construction and Application of Big Data System in Automobile Intelligent Factory

Yixin Yang\textsuperscript{1,2}, Jianjun Gao\textsuperscript{2}, Zhibo Pi\textsuperscript{2} and Konghui Guo\textsuperscript{1}

\textsuperscript{1} College of Mechanical and Vehicle Engineering, Hunan University, Changsha, China
\textsuperscript{2} Beijing Automobile Industrial Company Limited Zhuzhou Branch, Zhuzhou, China
Email: 13913840020@163.com

Abstract. Under the background of intelligent manufacturing, the application of industrial big data is the key to improve the utilization efficiency of manufacturing resources. Starting from the purpose of big data application and how to replace the efficiency of enterprise resource utilization, aiming at the construction of intelligent and transparent big data system in automobile factory, this paper combed the concept of big data system construction, discussed the construction ideas of big data system, constructed the big data system framework suitable for discrete manufacturing enterprises, and gave the methods of data acquisition, transmission, storage, processing and Analysis of the specific program. For big data applications, taking the enterprise intelligent operation cockpit, intelligent electronic process system and enterprise epidemic prevention and health cloud system as examples, the scene and application effect of big data application are demonstrated. The construction and application of big data system in intelligent and transparent automobile factory provides a reference model for further promoting intelligent manufacturing driven by big data.

1. Introduction

Big data is a technical system to effectively deal with the management and application of massive and complex data, and it also contains new thinking mode and innovative ability. From the perspective of strategic thinking, it is the key element of innovation and development in commerce, and even rise to the level of national strategy. The scientific community regards big data as the fourth paradigm \textsuperscript{[1]}, after experimental science, theoretical derivation and simulation. Big data is regarded as industrial oil in the new era by the enterprise community. To the enterprise community, who master the data, will master the future. From the perspective of complex scale manufacturing, information communication and sharing, and uncertainty system, the essence of intelligent manufacturing is to build a set of rules system for automatic data flow to resolve the uncertainty and optimize the allocation efficiency of manufacturing resources \textsuperscript{[2]}. The most important symbol to measure the intelligence level of an enterprise is not only the automation of production equipment, but also the automation of data flow according to set rules.

In the aspect of big data research and application, most of the achievements focus on the research methods of big data and the application of big data methods in certain field. For manufacturing enterprises, how to carry out the construction of big data system and how to implement its application, there is still no specific guidance and application cases \textsuperscript{[3]}. Most of the enterprises are aimless and have no idea how to start. Constructing enterprise big data system is not as simple as just buying a software package or building a database. To construct big data system should deeply understand the business of
the enterprise based on the actual situation, and integrate the factors of enterprise production equipment, software system, business process, etc. This paper took the intelligent automobile factory as example, analyzed and discussed the construction method and implementation path of big data system in manufacturing enterprises, and provided thinking and reference for enterprises to carry out the construction of big data system.

The construction of big data system first needs to consider its purpose and goal from the perspective of enterprise development and competition. The core competitiveness of enterprises is reflected in the efficiency of resource allocation. The core of resource allocation is scientific, efficient and accurate decision-making, which depends on the automatic flow of data. Secondly, it is necessary to consider how to realize data flow, integration and development from the perspective of smart manufacturing, so as to improve the capability and level of intelligent manufacturing. Finally, the implementation of landing, from the perspective of the enterprise, consider how to achieve each step from the aspects of data generation, collection, transmission, storage, processing, mining and application.

2. Architecture of Big Data System

Under the framework of intelligent manufacturing, the function of big data system is to establish a set of system that can make all of the enterprise data to be extracted, transmitted, integrated, stored and applied easily. So, how to build a big data system? Firstly, we need to clarify the relationship between intelligent manufacturing and big data system. Intelligent manufacturing is to realize the automatic and intelligent control of the manufacturing process by means of automation, digitalization, informatization, artificial intelligence, etc., so as to improve the utilization efficiency and manufacturing quality of manufacturing resources. Big data is the result of intelligent manufacturing on the one hand, on the other hand, integration, analysis, mining and reuse of big data will promote the development of intelligent manufacturing to a wider and deeper direction [4].

Next, the big data platform should be established to realize the centralized storage and interaction of data. For example, the big data platform represented by Hadoop supports data sorting, storage and calculation with 4V characteristics (large scale, multiple types, fast speed and low density) [5]. The most difficult part is to develop new applications based on big data platform. The enterprise big data team should not only be familiar with the business of the enterprise, but also have certain software development ability. Therefore, big data architecture should reduce the difficulty of development as much as possible.

Based on the above analysis, we have developed the big data architecture shown in figure 1. The framework gives the idea about how to promote the big data system quickly and effectively under the existing conditions, and clarifies the relationship among industrial Internet, data storage and calculation, traditional industrial software, data analysis and display port, industrial app, etc.

Data collection methods include sensors of automation equipment and facilities, intelligent equipment and information systems which can be used as data collection objects and main sources of data. Data transmission system includes Internet, Industrial Internet and Mobile Internet. Big data platform is responsible for induction, sorting and storage of massive data. It is different from traditional database management in that its scalable distributed storage system HDFS and parallel computing model and method MapReduce for large-scale data processing can adapt to the characteristics of big data 4V [6]. The application of data analysis, including data display and monitoring and various industrial app applications, refers to the application development based on the big data platform, rather than the application software system based on the traditional database.

The above architecture of big data system reflects the key elements and development focus known as the New Four Bases of Made in China 2025 [7]. The New Four Bases means Hardware (sensor and automatic control), Soft (Industrial APP), Network (Industrial Internet) and Platform (big data platform). The data acquisition and data transmission system corresponding to the Hardware and Network, and it needs to be gradually unified and upgraded. Platform is the big data platform that needs to be established and promoted quickly, and Software is the industrial app based on the big data platform.
3. Data Acquisition and Transmission

Sensor, PLC and intelligent terminal make the people, machines and services interconnected through industrial Internet, and realize horizontal, vertical and end-to-end data integration of different dimensions. Data collection and transmission are mainly carried out around the construction of CPS system. The following described the implementation and corresponding strategies of data collection and transmission from three aspects: mobile and PC side, production line equipment and the main data sources of existing IT systems.

3.1. Data Acquisition Though Mobile and PC Terminal.

Through the customized development of specific business app, data information acquisition can be realized though the mobile end and PC. The input information includes various forms and structures such as text, picture, audio, video, sensor information (such as GPS positioning, acceleration and other sensors) though code scanning and mobile terminal, etc. Information can also be imported into the database in batches through Excel and TXT files. As shown in figure 2, data is collected through the mobile terminal, uploaded to the big data cloud platform through the mobile Internet and the Internet, and data encryption and firewall can effectively ensure data security. With the promotion of 5G, mobile terminals will also expand from mobile phones and computers to cars, smart appliances, smart devices, etc., and the application scenarios will continue to expand.

3.2. Data Collection and Transmission of Production Line Equipment.

For the equipment, robot or execution unit on the production line, the control system is formed by IPC, PLC, HMI and industrial Internet. The engineering data in PLC, robot and other devices are not stored, so data collection of equipment needs to be achieved by deploying industrial collection software. Data collection must not affect the normal operation of the production line. Data for production equipment requires immediate (millisecond level), reliable (absolute security, usually double confirmation). Therefore, data transmission should be as concise as possible, reduce or avoid transmission through the industrial Internet, which consumes a lot of resources and causes industrial Internet congestion [8].

The data collection and transmission structure of production line equipment is shown in figure 3. The industrial Internet is the core of CPS space building. The following three points should be noticed in the system construction: (1) Use the new generation of industrial Internet as far as possible, such as PROFINET. Try to unify standards so as to avoid the difficulty of establishing communication between
production equipment. (2) Use Internet or mobile Internet as far as possible to alleviate the burden of industrial Internet. (3) Regional servers should be set up in equipment-intensive and data-intensive areas, and use Edge Computing technology to complete some data processing in the area. The data uploaded to the big data platform needs to be cleaned and sorted out in advance to improve the efficiency of transmission.

3.3. Data Collection and Transmission of Existing IT Information Systems.

This refers to the traditional database-based application systems including ERP, MES, CRM, PDM, LES, etc. The establishment of big data platform enables the data information of each software to be transferred or backed up to the big data platform. Enterprises can fully carry out data operation and application development on the big data platform according to their own requirement. This is equivalent to the interconnection of existing IT software through big data platforms, together with data from mobile and production line devices, to achieve wider data exchange and sharing. The analysis, monitoring and application of big data can be expanded quickly. At the same time, existing software can also generate new applications based on big data platform by developing and invoking data of big data platform. The development of Internet technology, especially the industrial Internet and the mobile Internet, makes it easier for information systems to migrate from the traditional architecture to the cloud architecture, as shown in figure 4. This kind of migration is the evolution of manufacturing resources from local optimization to global optimization, the expansion of enterprise cooperation from the inside of the enterprise to the industrial chain, and the upgrade of competition mode from single enterprise competition to ecosystem competition.

The above three modes of collection and transmission do not exist and operate independently, and often support each other to improve efficiency. We use the CPS system consisting of industrial Internet, mobile Internet and Internet+ to realize mutual recognition, real-time connection and effective communication of all relevant elements in manufacturing. The digitization and intelligence of R&D and design, process flow, production and manufacturing, product service and other processes restructure the elements and processes of production and manufacturing process. Many uncertainties faced by complex manufacturing systems have been made explicit, which makes it possible to build a new manufacturing system with automatic flow of technical data. The ways and means of optimizing the allocation of manufacturing resources are more abundant, convenient and efficient. For the transmission and control of device information, the CPS system is comparable to the central nervous system, and the big data system is comparable to the cerebral cortex, which stores and processes information from inside and outside, and then sends instructions to the executing agencies through the central nervous system.
4. Big Data Platform

Big data platforms are like data distribution centers, which are the storage, interaction, and computing centers of enterprise information. Big data platform enables enterprises to exchange and communicate all related data which can be digitized. It will inevitably improve the efficiency of information utilization and discovery. It will make all forms and sources of data have interactive platforms and make software application development easier. Unlike traditional IT systems, it is highly dependent on professional engineers for various data interface development.

To the big data platform, traditional IT giants such as IBM, ORACLE, HP, etc. launch their own integrated software and hardware solutions through mergers and acquisitions and development for enterprise use. Internet companies such as Google, Amazon, Ali, Baidu and Tencent form their own big data application circle and business model of the Internet based on their own application platforms and huge user groups and user information. The industrial big data field is represented by Siemens’ Mindsphere, GE’s Prex, and Kuka Robot’s KUKA Connect [9]. There are also many solutions for government, University and scientific research institutions.

Upon the promotion of big data application, many enterprises have introduced the outsourcing mode of IaaS, PaaS and SaaS. Outsourcing hardware facilities services such as servers, storage and networks through IaaS. Collaborative outsourcing of application development, such as virtual server applications, web application management, etc. is implemented through PaaS. Develop, manage, deploy and outsource system software services through SaaS. Whether an enterprise should build its own big data system or choose the above business outsourcing methods mainly focuses on data security, construction cost and sustainable development. For small enterprises, confidentiality is not strong and the development capacity is insufficient, public cloud platforms can be leased, and the above outsourcing mode can be selected as needed. For large and medium-sized enterprises or enterprises with strong professionalism and confidentiality, general APP cannot meet all the needs of enterprises, especially toward the unique management methods and processes of enterprises, it is recommended to build a private cloud platform and develop apps suitable for their own management.

The future big data architecture should be public cloud and private, which exist together and complement each other. Industrial APPs will grow like mobile APPs. There are many standardized and homogeneous general businesses in enterprises that try to use service outsourcing. Enterprises should devote massive work to software development involving core secrets, core competitiveness and unique business needs of the enterprise.

Based on the above, we chose Hadoop big data storage and computing system based on open source and sail soft data visualization system to build a big data platform, as shown in figure 5. Hadoop is an open source system that supports a variety of data sources and data types. Using Apache Hadoop distributed database, which is a leading enterprise data ecology framework, importing, integrating,
storing, analyzing, mining, external services from source data are all supported by corresponding components in Apache Hadoop ecology such as Map/Reduce, Hive and so on [10]. FineReport makes data display more intuitive and concise, reduces the difficulty of data analysis and processing, and makes it easier for non-IT professionals familiar with enterprise business to get started. The deployment of big data system follows the hierarchical structure of “data warehouse - data market (including processing) - data analysis application”, and uses the master-slave-app scheme to build a distributed architecture data platform.

Figure 5. The architecture of big data platform.

5. Big Data Analysis and Application

The core of big data application is to develop various industrial APPs on the big data platform to realize the application of big data analysis. The nature of the big data platform is “data + model = service”. Below are some specific cases to illustrate the development and promotion of big data applications.

5.1. Smart Cockpit

Smart cockpit is an enterprise operation monitoring and decision support system developed on a big data platform. As shown in figure 6, the enterprise cockpit displays information such as production, equipment operation status, logistics, quality, finance, etc. to the cockpit interface in real time and intuitively, which facilitates the monitoring and early warning of various information of the enterprise operation. Background data is a collection of data information for enterprises running on big data platforms. It shows the platform’s soft FineReport and FineBI designs integrated with Hadoop big data platforms without deep programming expertise. The cockpit supports both mobile and PC access, or electronic Kanban transported to production and business sites. The cockpit can be customized according to different management needs. The cockpit can be customized at company level, department level and team level, which makes it easy for different management units to grasp unit performance indicators in real time.

5.2. Intelligent Process System

This system achieves paperless and intelligent application of process files. As shown in figure 7, in the system, electronic process files are uploaded to the big data system, which enables related applications with production lines and different forms of query, supporting pictures, audio and video, word or excel, etc. We can query through the entrance of mobile enterprise WeChat, or through computers, iPads and line-edge smart terminals. It supports flexible query methods such as operation number, operation name, part name, two-dimensional code scanning, etc. Passwords are not required for intranet or WiFi queries, passwords and personal authentication are required for external networks, and the pages displayed all have confidential letters and watermarks for querying personal information. If exported, traceability can be achieved. Intelligent process system and production line equipment are connected through the industrial internet. The factory body and the whole vehicle are connected with RFID chips. The RFID identification system and the industrial Internet are passed to the line edge device or process terminal.
The line edge device invokes the process from the big data platform according to the product information, chooses the corresponding process parameters, shows the operation essentials and precautions of the position. Through the intelligent process system, the correlation between process parameters and line edge equipment and products can be achieved, which lays a foundation for the future customized production of automobiles.

Figure 6. Enterprise operation cockpit (Notes: 智能透明工厂运营驾驶舱-Intelligent transparent factory operation cockpit; 能源管理-Energy management; 全员经营-All staff operating; 质量管理-Quality assurance).

Figure 7. Intelligent process system.

5.3. Employee Healthy Cloud System
During the novel coronavirus pneumonia epidemic prevention, Healthy Cloud System is required to complete development and application based on the big data platform within just three days, which verifies the efficiency of big data application development based on the big data platform definitely. According to the requirements in epidemic prevention, the system can statistics company employees’ daily health status, travel, residence, going out during work, close contact and other information. Once the epidemic situation occurs, the information of close contacts can be quickly transferred out. The system is developed on the basis of big data platform and FineReport software. Mobile data input and submission ports are designed through Enterprise WeChat, including employee’s work number, employee's name, home address, office, dormitory, temperature, displacement, outgoing, close contact information, etc. Location information is read automatically through mobile phone GPS positioning, and other information is entered manually. To improve the efficiency of input, and automatically read the latest information in the database history for information that changes frequently, you can modify it according to the current information and set a temperature input error reminder (the range of allowed...
input is 35-42 degrees). Employees report daily according to WeChat reminders, aggregate data into big data platforms, and use FineReport in the background to customize different reports for statistics, queries and display, as shown in figure 8. The functions of approval, attendance and reporting of abnormal information on employees’ going out are integrated into the system at the same time, which strongly supports the improvement of management efficiency during epidemic prevention. The effective operation of this system has played a good role in demonstrating the application of big data expansion.

**Figure 8. Employee health cloud system.**

According to the above ideas, the company has developed a stop management system, parts inspection system, human resource performance system, welding parameters monitoring system, body accuracy system, process parameters monitoring system, quality information system based on the big data platform. The application development based on big data platform has stimulated the enthusiasm of team innovation, and has played a significant role in promoting the efficiency of the company's resource utilization.

6. Conclusion

Big data technology enables data flow to produce value and enhance the competitiveness of enterprises through value creation. By promoting the interactive innovation of big data and business management processes and modes, technology progress, management change and business process optimization are achieved, and new capabilities such as enterprise personalized customization, lean management, risk control, supply chain collaboration and market quick response are formed. Help enterprises restructure the organizational model of production management services to form sustainable and differentiated competitive advantage.

The application of big data provides more accurate statistical analysis basis and target for the improvement of production efficiency, wider technical methods and means for product quality improvement and technical improvement, clearer basis and direction for the improvement of company cost control and asset utilization efficiency, and more accurate decision-making basis for company operation management. The construction and application of big data system has promoted the level of Intelligent Manufacturing of the company, promoted the establishment of the full participation business model.

Zhuzhou Brunch was evaluated as the management accounting pilot demonstration enterprise by Ministry of industry and information technology of China, and was honored Auto Intelligent Manufacturing Demonstration Factory by Hunan province. The big data system will continue to improve and promote the level of business management and competitiveness of enterprises with the popularization of 5G, edge computing and other new technologies’ development.

Acknowledgments

Authors wishing to acknowledge assistance of colleagues from BAIC MOTOR Group Zhuzhou Brunch.
Beijing Redoop System Technology Co., Ltd supported the construction of big data system from providing software through cooperation. The authors would like to thank Mr LUO Guobiao and Miss ZHOU Yan for valuable technical discussions.

References
[1] Hey T, Tansley S and Tolle K 2009 The fourth paradigm: Data-intensive scientific discovery Proceedings of the IEEE 99 (8) 1334-1337.
[2] Guo Q and Zhang M 2009 A novel approach for multi-agent-based intelligent manufacturing system Information Sciences 179 (18) 3079-3090.
[3] Wang Q, Li J, Chen X, et al. 2016 Summary of big data analysis platform construction and application Integrated Technology 5 (24) 4-20.
[4] Jagadish H V, Gehrke J, Labrinidis A, et al. 2014 Big data and its technical challenges Communications of the ACM 57 (7) 86+88-94.
[5] Gong X, Li B, Chai X, et al. 2014 Technology review of big data platform Journal of Systems Simulation 26 (3) 489-496.
[6] Lynch C 2008 Big data: How do your data grow? Nature 455 (7209) 28-29.
[7] An X 2019 Logic of Reconfiguration-Digital Transformation (Electronic Industry Press) Chapter 3 pp 37-40.
[8] Yang Y, Yuan Z, Pi Z, et al. 2018 Construction and implementation of an intelligent and transparent automobile factory China Mechanical Engineering 29 (23) 99-106.
[9] Liu Q and Qin S J 2016 Perspectives on big data modeling of process industries Acta Automatica Sinica 42 (02) 3-13.
[10] Tian X, Zhou Y, Bi Z, et al. 2019 Distributed computing and storage technology based on Hadoop architecture and its application Journal of Shanghai Electric Power Institute (1) 73-77.