ISSN 2959-6157

# Advantages of Big Data Analysis and AI Technology in Data Collecting And Processing of Preventive Maintenance

**Xinyuan Liu**

The High School Attached to Hunan Normal University,Changsha,410000,China;

**Abstract:**

In recent years, the emergence of Cyber-Physical systems (CPS), big data, cloud computing, and industrial wireless networks have promoted the development of Industry 4.0. Intelligent manufacturing systems are the leaders of this change. In smart manufacturing, the emphasis will be placed on using technologies such as big data analytics, cloud computing, edge computing, and artificial intelligence (AI). The advantages of new technologies are significantly reflected in the comparison of preventive maintenance and traditional maintenance based on intelligent technology: preventive maintenance significantly reduces the cost of prevention and, at the same time, dramatically improves the accuracy and timeliness of maintenance, bringing inestimable value to intelligent manufacturing. This paper will analyze new technologies' innovative advantages and technical prospects through big data analysis and specific application scenarios of AI technology in preventive maintenance. Firstly, in data acquisition, this paper examines the architecture of preventive maintenance systems based on big data. Then, it analyzes the composition of the data pipeline: in the realization of data collection and processing in the field of cloud computing and edge computing, and the aspect of data processing, it discusses the participation of AI technology in data understanding and processing and the unprecedented changes it brings.

**Keywords:** Preventive Maintenance; Big Data; AI technology.

## 1. Introduction

In recent years, advances in information physical systems (CPS), big data, cloud computing, and industrial wireless networks have driven the implementation of Industry 4.0 [1]. Information technology is infiltrating all aspects of manufacturing systems [2] as well as other fields [3], accelerating the generation of big data in manufacturing [4]. Big data in manufacturing can be divided into three main types: device data, product data, and command data. Device data refers to the information generated by machines and equipment in manufacturing, such as operating temperature, speed, and power consumption. Product data refers to detailed information about the products being produced, including material composition, size, and quality control indicators. And about the command data refers to the instruction and control signals sent to the manufacturing equipment. It is the director of the production process. New applications in manufacturing as well as new solutions are drawing inspiration from big data analytics, and new manufacturing process management and processing models will bring unprecedented changes. Manufacturers, as monitors of the process, also get transformative experience improvements from it. They collect, store, and analyze the massive data of the manufacturing process through the cloud platform, and obtain comprehensive accurate, and even visual data monitoring. Beyond the examples provided above, big data can also enhance various aspects of manufacturing and even provide creative applications. This was previously unimaginable. Similarly, in terms of equipment maintenance, which is the focus of this paper, proactive preventive equipment maintenance based on big data is a typical example. In this new system, large amounts of production data need to be collected, and those data are studied to predict potential equipment failures and schedule maintenance activities accordingly. This inventive new maintenance structure significantly reduces the downtime required for maintenance and increases the overall efficiency of the manufacturing process. Proactive equipment maintenance fully embodies the evolution brought by big data analytics in the manufacturing field discussed earlier.

Notably, equipment maintenance plays an important role in intelligent manufacturing, directly affecting equipment's service life and production efficiency [5]. Preventive maintenance, in particular, has received a big deal of attention due to its ability to predict and prevent equipment failures before they occur. Its preventive capabilities need some help: RFID and the Internet of Things (IoT). With these techniques, preventive maintenance can track the life cycle of a product and establish important information links with manufacturers. Under this situation, the emergence of “big data” allows for a new generation of maintenance, such as preventive and predictive maintenance. Representative preventive and predictive proactive maintenance has apparent advantages over traditional decentralized maintenance by significantly improving the efficiency of equipment maintenance. Traditional methods usually focus on the logical relationship between equipment, shop floor, and plant during operation and maintenance. However, active maintenance simplifies the operational logic by mapping each element directly to the corresponding maintenance resource, Thus breaking this logical dependency. In addition, proactive maintenance enables dynamic monitoring of the status of the entire plant through a unified visual management. This monitoring approach improves the timeliness and accuracy of maintenance compared to traditional decentralized maintenance practices with after-the-fact reporting modes while providing more precise insight into equipment operating status for more informed decisions. Real-time proactive maintenance offers significant advantages in terms of response time compared to traditional decentralized maintenance and post hoc reporting models. Traditional maintenance models often employ a layered fault reporting approach, which often leads to delays in transmitting and processing fault information. Real-time active maintenance Through active reporting and real-time response, faults can be discovered and resolved promptly, improving the efficiency and flexibility of maintenance operations. That is to say, as maintenance gradually moves in this direction, the line between utility and repair will no longer be clear, and “maintenance will happen as long as the product is in use.”[6]

Through the research of big data technology and the achievements of current researchers, this paper analyzes the distinctions between preventive maintenance and traditional maintenance in many aspects such as working structure and efficiency. Meanwhile, this paper investigates the application of big data technology in the current field and studies the development prospects of new technologies including AI in manufacturing. Based on the aforementioned research work, this paper is divided into five parts. The rest of this article follows. Section II covers data acquisition, big data, and cloud computing in manufacturing. Section III gives examples of data pipeline architecture and cloud computing. Section IV presents the AI technology method for data processing. Section V summarizes this thesis.

## 2. System Architecture

In the traditional maintenance model, we usually follow a three-tier structure: production line maintenance, shop floor maintenance, and factory maintenance. In this model, however, when there is a malfunction or problem, the staff generates non-real-time reports, passes layer by layer, and solves the problem in time. However, this model has one obvious drawback: lack of predictability. We can often only act accordingly after a problem arises, which inevitably increases production costs and complicates maintenance procedures.

The proactive preventive maintenance model based on big data has completely changed the situation. In this model, we can collect product data, equipment status updates, facility logs, equipment alarms, production process information, and other relevant data through industrial wireless networks. These real-time data are transmitted to the cloud, where correlation analysis is carried out to reveal various relationships within the equipment. In this way, we can spot potential problems in time and predict possible equipment failures, enabling us to implement necessary maintenance measures in advance proactively. The significant advantages of a proactive preventive maintenance model based on big data compared to the traditional maintenance model can be concretely reflected in the technical aspects of data acquisition and data analysis. First of all, in terms of data collection, we use the industrial wireless network to achieve real-time data acquisition which eliminates the limitations of traditional inspection and manual recording: and reduces the error caused by human factors. At the same time, real-time data collection can also improve the accuracy and efficiency of data collection. Second, we leverage new technologies in data processing: big data analytics and machine learning techniques. These new technologies can deeply process and study big data, and discover the rules and problems hidden behind the data [5], letting data processing no longer require inefficient ways like human detection based on experience.

In this application context, we learned from the results of the researchers that they proposed the service-oriented industrial application architecture OPCUA, which is considered as one of the most promising breakthroughs in the field of data integration. The OPCUA architecture enables information sharing by integrating data between industrial systems, providing enterprises with an understandable data view. It can not only effectively solve the problem of data silos but also improve the accuracy and efficiency of decision-making through unified data management. Therefore, OPCUA architecture has been widely used and recognized in industrial automation, intelligent manufacturing, and other fields. Active maintenance technology based on big data is essential to OPCUA architecture. The key idea is to carry out preventive maintenance on industrial equipment in real-time or offline to improve the reliability and working life of equipment. Active maintenance technology mainly includes two types: mechanism of real-time active maintenance (MRAM) and mechanism of offline Prediction and analysis (MOPA). The former meets the high real-time requirements of equipment operation and maintenance. MRAM aims to discover errors and handle real-time alarms by collecting and analyzing equipment data. This mechanism can ensure that the industrial system always maintains a stable working state during the production process and avoids production interruption or quality decline due to equipment failure. The latter mines the potential disturbance of maintenance projects, while MOPA focuses on the potential risk mining and failure prediction of maintenance projects. Through the analysis and mining of historical data, the potential problems and hidden dangers existing in the operation of equipment are found to provide accurate maintenance suggestions and preventive measures for maintenance personnel. In addition, MOPA is responsible for proactively implementing the required maintenance measures to ensure the equipment's regular operation and production efficiency. In practical applications, MOPA can forewarn potential failure risks in advance, providing sufficient time for businesses to carry out maintenance and repairs. This reduces the losses caused by equipment failures and improves enterprises' production efficiency and economic benefits (Figure 1).

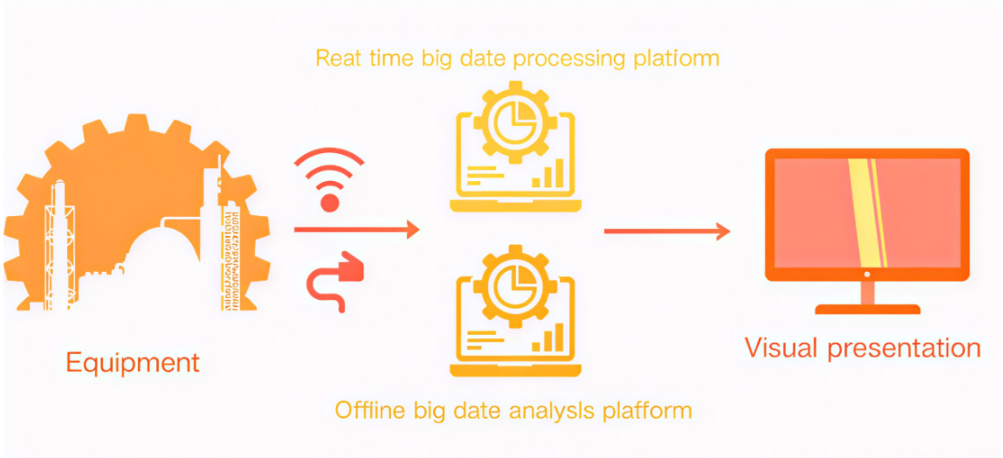


**Figure 1. The new mechanisms: MRAM and MOPA process equipment data to fix errors in real-time and predict errors before they occur. Then, equipment information will be presented on visual platforms to help producers understand the state of all equipment.**

The critical components to implementing MRAM and MOPA are cloud layers, which play a crucial role in data processing and mining. To ensure efficient real-time processing and data mining, the data processing system must have highly parallel processing capabilities, high throughput data transmission, and robust communication. In the current research on real-time data processing, researchers have discussed the application scenarios of STORM real-time computing system and Hadoop distributed parallel batch processing system [5]. We can see that the construction of data storage, monitoring, and computing modules can be realized by the integration of these two systems. Meanwhile, the efficiency and accuracy of real-time data processing can also be significantly improved in this integration. In addition, the researchers verified the effectiveness of the offline prediction algorithm, a technique that is critical to MOPA, by conducting a series of experiments. Taking a machining center as an example, researchers compared the influence of offline prediction algorithms and traditional empirical estimation algorithms on tool life. By comparing their application in traditional machining processes such as cutting and drilling, they observed that the offline prediction algorithm provides more accurate and effective tool life prediction.

Therefore, we can conclude that the OPCUA architecture and its proactive maintenance technology based on big data provide strong support for data integration and preventive maintenance in the industrial field. By integrating data resources between different systems, OPCUA realizes information sharing, providing enterprises with clear data views and decision support. Proactive maintenance technology based on big data performs preventive maintenance on equipment in a real-time, offline way, improving the reliability and service life of equipment. With the continuous progress of technology and the continuous expansion of application fields, it is believed that OPCUA and its related technologies will have a wider application in future industrial development.

In addition, it is noticeable that data format and data preprocessing are also important links in the construction of a proactive preventive maintenance system based on big data. Since the types and formats of data generated by different devices may be different, we need to carry out unified format conversion and preprocessing of data to ensure its accuracy and consistency.

## 3. Data Pipeline

3 basic parts are included in the data pipeline wherein a modern manufacturing environment: data collection, data processing, and data analysis using AI. This new structure benefits the efficient use of more data in an industrial setting, and enables manufacturers to gain valuable insights into their operations. (Figure 2)

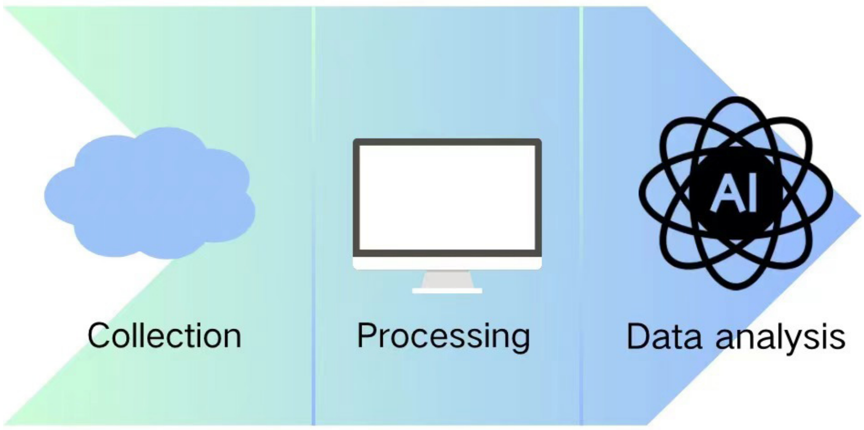


**Figure 2. The data pipeline of modern manufacturing includes three parts: data collection, data processing, and data analysis.**

From Figure 2, the data collection is the fundamental of the next analysis for preventive maintenance. In data collection, we utilize industrial networks or the industrial Internet of Things (IoT) to collect and transmit various types of data to computing servers. These data include environmental indicators. After the data collection step, manufacturers can learn the production process. In this new pipeline, we used AI techniques and integrated them into the data mining and analysis phases. This innovation further increases the value of the pipeline. The ability of AI algorithms to sit through lots of data not only identifies patterns and trends that human analysts might miss, but also enables manufacturers to understand production efficiency. As data pipelines are combined with new technologies, manufacturers can make more informed decisions, optimize operations, and reduce costs.

(1) Data Collection: Edge and Cloud Computing Paradigms

When we review some examples derived from the real world, we can figure out cloud services hosted and provided by companies e.g. Amazon, Apple, Google, Microsoft, and Facebook that give manufacturers access to powerful computing resources on a pay-as-you-go basis. It is generally accepted that the intelligence and resource capabilities required for IoT data processing are concentrated in cloud data centers. However, as technology advances and application scenarios become increasingly complex, this notion is being challenged. Traditional cloud-centric approaches to IoT, such as Amazon IoT and Google Cloud Data Stream, are shifting to a more distributed model. This shift is designed to take full advantage of intelligent and programmable cloud services at the edge of the network, Including intelligent gateways (e.g. Raspberry Pi3, UDOO board, ESP8266) and network function virtualization solutions (e.g. Cisco lOx, HP OpenFlow, Middle-box Technologies). These edge data centers offer computing and storage capabilities on a smaller scale than traditional cloud data centers while playing a vital role in real-time data processing. Two main benefits can be realized by bringing IoT data processing activities closer to the source or receiving point of the data transmission. First, it helps conserve energy consumption in resource-constrained edge devices. Under the current resource management model, where these devices are constantly uploading data to cloud centers to process tasks, by moving some of these tasks to edge data centers, we can both reduce the amount of data that is unnecessarily transferred and reduce the energy consumption of the devices, significantly extending the life of the devices while improving overall energy efficiency. Secondly, this distributed processing model can also minimize unnecessary network bandwidth consumption. In the centralized cloud data center model, large amounts of data need to be migrated between the Internet and public/private data centers. This not only exacerbates the risk of network congestion but also introduces potential delays and loss in data transmission. There is a solution: by distributing the data processing tasks over multiple nodes at the edge of the network, the amount of data transmitted through a single node can be reduced, thus reducing the network bandwidth consumption[7]. This distributed processing reduces communication latency and dependence on migrating large amounts of data across the Internet and public/private data centers. Therefore, as an extension of existing IoT devices, edge data centers enable real-time monitoring and optimization of manufacturing processes by providing improved processing and storage capabilities. At the same time, edge data centers can also adopt various mechanisms to process data on behalf of IoT devices. It is only migrated to a remote cloud data center when more complex analytics are required or edge processing capabilities are not available. As the number of IoT devices continues to proliferate, the demand for efficient processing and storage of the data sets they generate is also increasing. We can see that this fragmented shift has a profound impact on IoT applications. Edge data centers meet these requirements well while improving the performance and reliability of IoT applications. At the same time, its functionality and performance are also constantly being upgraded with the advancement of technology. We can expect to see more intelligent algorithms and machine-learning techniques applied to edge data centers for more efficient data processing and analysis. Meanwhile, as network technology evolves, collaboration and communication between edge data centers will also become more efficient and reliable.

(2) Data Processing: AI Technology

The use of AI technology and intense learning has become the most effective and cutting-edge method in fault identification of systems and equipment based on big data analysis. The importance of AI in extracting valuable knowledge from big manufacturing data cannot be overlooked, as it forms the key to enabling intelligence in industrial environments. The Data-Information-Knowledge-Wisdom (DIKW) hierarchy, often referred to as the DIKW pyramid, provides a comprehensive framework for understanding the evolution of Data into intelligence. (Figure.3) This hierarchy includes four layers, with data as the foundation. Information is the next layer, representing organized and processed data. Knowledge is on the top of information and involves understanding and interpreting information, often through patterns and insights. Finally, wisdom, located at the top of the pyramid, means the application of knowledge to making informed decisions and judgments.

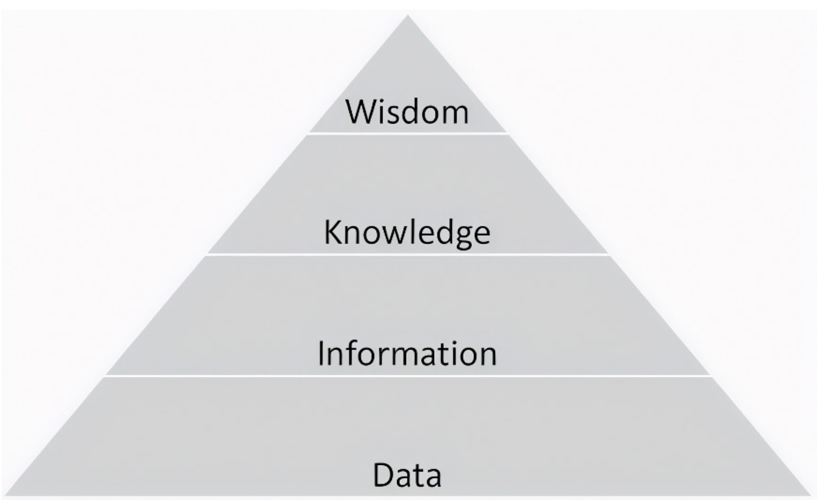


**Figure 3. DIKW hierarchy, also known as the DIKW pyramid, consists of four parts and has an ultimate goal: getting wisdom.**

The DIKW hierarchy provides a valuable framework for understanding the transformation of raw data into valuable insights and wisdom when leveraging AI technologies. Not only as the embodiment of artificial intelligence for data research, DIKW hierarchy is widely used in the field of information science and information management as a data research tool. In research in information science, researchers use the DIKW model to explain the logical and conceptual frameworks that they find relevant. Especially in the study of those frameworks related to knowledge and epistemology, the role of this model is particularly significant. In another area, business managers responsible for information management have recognized the importance of the DIKW model in solving real-world challenges, especially those involving information utilization and management[8].

## 4. Discussion

In the previous discussion in this paper, we can learn a completely new data analysis system: a manufacturing maintenance knowledge base built on manufacturing big data. Its functions include: information analysis, which enables manufacturers to understand the complex workings of their equipment; The identification of patterns and trends that allow and predict potential problems before they occur. By building this knowledge base, manufacturers can develop proactive and predictive strategies that significantly reduce downtime. It relies on a combination of edge computing and cloud servers in the data collection phase to process the large amount of data generated by manufacturing devices. By using an onboard server near the device, it allows data processing to take place online and near the data source. As a result, it can provide low-latency (real-time) services. This method can quickly respond to the change of equipment state, and is especially suitable for real-time monitoring applications. At the same time, we can not ignore the high computing power of cloud servers, which makes them useful in processing big data offline. Cloud environments provide the scalability and flexibility needed to handle the growing volume of data generated by manufacturing operations. In addition, cloud environments enable manufacturers to share data and collaborate with other organizations, thus facilitating innovation and knowledge sharing within the industry. And in the next data processing stage, it is crucial to eliminate redundant and misleading data. This cleaning process ensures that only relevant and accurate information can be used for further analysis. As the data is cleaned, the program abstracts it into real-time or historical big data analytics. This analysis can be used as a basis for equipment maintenance decisions and can also provide valuable insights into the status and performance of manufacturing equipment. Therefore, data processing devices such as computer servers play an important role in data processing. The computer server performs the data analysis results are then transmitted to a data management or visualization system. So these systems allow manufacturers to visualize the data meaningfully, such as through charts, graphs, and dashboards. This visualization not only makes the data more accessible to understand but also enables manufacturers to quickly identify trends and patterns. In addition, it facilitates decision-making, enabling manufacturers to make correct choices in equipment maintenance and operational strategies.

In the future, we can tell that the development of various information transmission processing equipment, such as computer servers, will also drive the further improvement of proactive preventive maintenance capabilities and effectiveness, bringing more opportunities for application and providing more timely, accurate, and economical equipment maintenance for the manufacturing industry.

## 5. Conclusion

In conjunction with the development trend of digital manufacturing in the era of Industry 4.0, this paper discusses the advantages of big data analysis and artificial intelligence technology in proactive preventive maintenance, comparing it with the traditional mode and the technological innovation of data pipelines in the new era. First, within the framework of the preventive maintenance architecture, new intelligent maintenance methods utilizing big data analytics show high efficiency and predictability. This intelligent maintenance method not only significantly reduces maintenance costs but also enhances the flexibility and adaptability of the entire maintenance process. Second, in terms of data transmission, big data-driven maintenance has significant advantages with the help of IoT technology. With the further development of Industry 4.0, big data and artificial intelligence technologies will play an increasingly important role in preventive maintenance. In the future, we can foresee that these technologies will further promote the intelligent and automated development of preventive maintenance, creating more excellent value for enterprises. Besides, we also need to pay attention to the equipment and algorithms that support this intelligent maintenance and believe that the development of these technologies will also bring more powerful utility and more application possibilities for preventive maintenance.

## References

[1] S. Wang, J. Wan, D. Zhang, D. Li and C. Zhang, “Towards the Smart Factory for Industry 4.0: A Self-organized Multi-agent System Assisted with Big Data Based Feedback and Coordination,” Computer Networks, vol. 101, pp. 158-168, 2016.

[2] S. Jeschke, C. Brecher, H. Song, and D. Rawat, “Industrial Internet of Things: Foundations, Principles and Applications,” Springer, Cham, Switzerland, pp. 1-715, 2017.

[3] Z. Lv, H. Song, P. Basanta-Val, A. Steed, M. Jo, “Next-Generation Big Data Analytics: State of the Art, Challenges, and Future Research Topics,” IEEE Transactions on Industrial Informatics, no.99, pp.1-1, 2017.

[4] Y. Sun, H. Song, A. J. Jara and R. Bie, “ Internet of things and big data analytics for smart and connected communities,” IEEE Access, vol. 4, pp. 766-773, 2016.

[5] J. Wan, S. Tang, D. Li, S. Wang, C. Liu, H. Abbas, and A. V. Vasilakos, “A manufacturing big data solution for active preventive maintenance,” IEEE Transactions on Industrial

Informatics, vol. 13, no. 4, pp. 2039– 2047, 2017.
[6] R. Y. Zhong, S. Lan, C. Xu, Q. Dai, and G. Q. Huang, “Visualization of RFID-enabled shopfloor logistics Big Data in Cloud Manufacturing,” The International Journal of Advanced Manufacturing Technology, vol. 84, no. 1-4, pp. 5–16, 2016.
[7] D. Georgakopoulos, P. P. Jayaraman, M. Fazia, M. Villari, and R. Ranjan, “Internet of things and edge cloud computing roadmap for manufacturing,” IEEE Cloud Computing, vol. 3, no. 4, pp. 66–73, 2016.
[8] M. H. Fricke,´ Data-Information-Knowledge-Wisdom (DIKW) Pyramid, Framework, Continuum. Cham: Springer International Publishing, pp. 1–4, 2018.