

# Key Technology and Application Analysis of Integrated Big Data Platform of Smart Power Plant

Changjun Yu

Dynamics Engineering Department, Anhui Electrical Engineering Professional Technique College, Hefei 230051, China

**Keywords:** Data, Integrated Platform, Power Plant

**Abstract:** in the Future, the Development of Information Technology in the Power Industry is Inseparable from the Construction of Smart Power Plants. the Core of Smart Power Plant Construction is to Establish an Integrated Big Data Platform in the Electric Field to Realize Intelligent Processing of Various Information in the Electric Field. by Analyzing the Development Trend of the Power Industry, It Can Be Found That Data Information and Data Sources Occupy a Large Proportion of the Information Resources of the Power Industry. Only an Integrated Big Data Platform Can Be Established, and Efficient Data Processing and Calculation Can Be Realized to Ensure the Accuracy of Data Information. . in the Process of Practice, This Integrated Platform Can Provide Many Conveniences for Enterprise Data Processing.

## 1. Introduction

The core of “smart power plant” is “wisdom”, which is mainly to let the various devices in the system have the ability to self-run, self-manage and self-diagnosis. The key is to realize the sorting and calculation of data. In recent years, with the improvement of science and technology, various mechanized management technologies applied in electric fields have been increasing, such as distributed control systems and information management systems that can realize digital control, with the widespread application of digital technology in the power industry. The main value of data resources is becoming more and more obvious. In recent years, with the development of the power industry, the growth trend of data resources is exploding. Various data in the power industry are coming together. The data types are complex, with wide distribution, non-deterministic, and multi-source heterogeneity. In the process of data collection and storage, it is also necessary for the staff to accurately extract, process and extract the knowledge through practical experience, but this is a certain way just like this will only accept the development in the right direction. It is an important direction of intelligent electric field research to obtain power station operation and maintenance information through intelligent means, to achieve efficient monitoring of the running equipment in the power station, and to find faulty equipment from many operating equipments at the first time. The electric field equipment will generate a large amount of data information during the operation and maintenance of the power station. The traditional methods can not meet the needs of the current stage of electric field development. Therefore, it is imperative to apply a more efficient and stable method. Next, the data processing method in the data platform integrated by the smart power plant is deeply discussed, and the means to stabilize the operation of the electric field is proposed, and the core technology in the intelligent power plant is analyzed.

## 2. The Construction of an Integrated Data Platform

### 2.1 Source of Data in Smart Power Plants

At this stage, the information level of electric fields is continuously improved, the sensing capacity of operating equipment is continuously enhanced, and the management mode is further innovated at the traditional level. In the process of self-operation, there are a lot of production data and operation management data that need to be processed. [1].

Production data accounts for a large proportion in the intelligent electric field big data organization. The data has great practical value and analytical value, mainly from the actual parameters in the operation of the equipment, the output and output time of the production products, and the equipment failure. The situation, etc., these are important data that appear in production logs and reports from time to time, and can be optimized and improved at the level of key information extraction and equipment maintenance. Power plants need to have a large amount of information data as the basis for realizing automation operations, so some professional and multi-angle data information is generated, which is called operation management data. When carrying out operational management data processing, the effective information contained in it can be presented in an optimal form to provide effective decision-making for the operation and management of the power plant.

According to the type of data structure, smart power plant big data can also be divided into structured data and unstructured data. The so-called structured data is in the process of normal operation of some fixed operation equipment such as generators, boilers, etc. Actual data generated, device information, and so on. Unstructured data mainly refers to materials that are auxiliary to the production process, such as text, images, audio, and text, which are biased toward technology.

## **2.2 The Structural Framework of the Integrated Data Platform**

The level of data generated in smart power plants needs to be very extensive. A large amount of information data will emerge from time to time. In order to make the data processing work efficient and stable, it is necessary to build an integrated big data platform. [2].

In the integrated data platform, the quality of the equipment is monitored by pre-processing the initial data of the power plant, and the effective data is put into the database. In the detailed analysis of data, the data in the database is generally used. An early warning system exists in the database, and the data model is effectively analyzed and processed. The data mining system can realize the in-depth mining of information resources in the database. The comprehensive technology platform provides assistance for users to view data and conduct personal analysis and statistics. The big data platform is organized in the form of a layered architecture to facilitate the expansion of various services in the system, to continuously integrate the business requirements in the system, to find the most shortcut way to formulate new task modules, propose new functions, and adapt to the electric field. The changing needs, improve work efficiency, improve development speed in the future work development process, reduce the difficulty in the development process, and reduce the running cost [3]. In the smart power plant integrated big data platform architecture can be divided into the following levels:

1) data information acquisition layer. It mainly realizes the collection and storage of data information in the field, and can directly collect data information from multiple databases, and connect the power plants as a whole. For example, when collecting data information from a real-time database, it can support multiple acquisition methods.

2) Data information delivery and processing layer. The data information is transmitted from the collection port into the database and becomes a data stream. After the data is authenticated, it can be directly written into the database. Data authentication needs to apply multiple data calculation models to achieve multiple data calculations, verify that the data generates errors, and promptly issue warnings and high red marks for the error data in the database, and replace or delete the data according to the actual situation. The accuracy of the data information is guaranteed.

3) data storage layer. The data storage layer is the core of data information storage and management. When collecting and calculating data information and processing, the data information is first stored together, so that all work is carried out uniformly. The data processing center includes two parts: data storage processing and data disaster recovery processing. The two are compatible with each other, remind each other and function independently.

4) business processing layer. In the business processing layer, it is organized and planned according to the business requirements of different operating modules in the power plant, integrated and processed according to the contents of different data in the sector, and the processed data results

are stored in the data storage center, and then the next step is integrated. Analysis and data mining work.

### **2.3 Features in the Integrated Big Data Platform**

The integrated big data platform can realize the discovery, analysis and resolution of problems, and has a targeted management plan. Use real-time monitoring to find the location of problems in the power plant, find out the cause of the problem through comprehensive analysis and diagnosis, and develop guiding problem solutions and problem management mechanisms. Based on the personal experience of the staff and the theoretical basis of professional knowledge, the problem is managed and analyzed, and then the smart flow plant integrated big data platform business process table and platform business data flow table are developed [4]. The process mainly includes the following five parts:

1) basic data. There are three main steps in the process of sorting the basic data. First, real-time information is collected from the power plant operating equipment, and then manually entered into the database. Finally, the information with the use value is analyzed and collected for use by other systems. The collected information resources are very extensive, such as document information, device operation data, real-time monitoring data, etc. The information content may also be real-time data with utility value obtained from other sites.

2) data identification. The data identification mainly judges the use value and accuracy of the data information by formulating the data, and promptly warns the error information existing in the database, and the data analyst changes or deletes the data according to the actual situation.

3) Real-time analysis. Real-time analysis mainly includes real-time calculation and real-time alarm calculation. The former pays more attention to computing tasks, while the latter pays more attention to the entry process. It mainly develops targeted service programs based on the needs of business capabilities, and completes the analysis and processing of data.

4) statistical analysis. Statistical analysis is also performed by means of statistical calculations. Because of the performance requirements in the system, in the first two process data identification and substantial analysis data calculation, the accuracy of the data cannot be guaranteed. Therefore, the authenticity of the data is verified by statistical analysis. It is an inevitable requirement for the process development, and mainly verifies and checks the data identification results and the real-time analysis results of the data.

5) big data analysis. This step provides data reference for the development of various important decisions in the power plant, and combines the big data analysis platform to find the most suitable mining algorithm, realize the calculation and real-time analysis and processing of the basic data, and realize the high-value data by formulating the form of the decision-making scheme. use efficiently. Display the content of the information using visual tools as an important support information in decision making.

## **3. Key Technologies Applied in the Integrated Big Data Platform**

### **3.1 Realize Multi-Source Heterogeneous Data Processing and Storage**

In the structural composition of smart power plants, the extraction efficiency of high-value information data is much higher than that of traditional data management systems. Due to the wide range of data sources and diverse formats, data resources have always been complex and diverse. By building an integrated big data platform, the following functions can be realized:

1) can quickly filter out the efficient use of data from the database, which is an important function to improve on the basis of decision support, according to user needs, can immediately sort out all the data generated in the history of a device [5]. It provides convenient conditions for enterprises to collect different systems and equipment data, and can analyze the operation trend of equipment and discover the hidden content in the data.

2) Sort data according to chronological order. This function has many advantages over the traditional key value data record form. It can query data according to the order of large data sets or

directly search for relevant time segments. Provide convenient conditions in the process of real-time data recording, accelerate the improvement of data read and write performance, and efficiently identify data acquisition data in the data system. In this way, the resolution of the data information is significantly improved, and the improvement is continuously promoted.

3) can connect the data information processing process with the data source, accelerate the efficiency of data information processing, obtain information directly from the running device, and quickly integrate into the enterprise data processing center to integrate the associated data for compression. , to achieve efficient and rapid storage, greatly reducing the difficulty of information regeneration.

### **3.2 Establish a Modified Model of Missing Points and Noise Points of Time Series Data**

In the integrated data platform, the time series model can be used to efficiently identify various operating states of the device, sort in the form of time series, detect and identify the running of the data, and analyze whether the data is abnormal according to the presentation state of the data, thereby Perform an overall analysis of the device that generates the abnormal data to determine whether it is a device failure or useless data due to improper operation [6]. If it is found that the device is running during the detection process, then the time-sequence intervention model is needed to fit and extract the fault information. When data is cleaned, adjustments and changes can be made according to the state of the outliers, thereby completing the task of correcting the noise point data and filling in the missing values. Compared with the traditional method of correcting noise points, this method of eliminating unnecessary data can perform efficient correction of data without noise points and missing value data. Reduce the loss of data due to improper data replacement, and make the dynamic change of time sequence always present.

## **4. Conclusion**

Through this article, we introduce the structural framework of the smart power plant integrated big data platform, and develop an efficient and secure data processing platform based on the characteristics of data processing in the power industry. In the future development of the database, it is also necessary to make reasonable planning of its internal system, reduce the coupling between modules, prevent data loss and leakage problems, and provide convenient conditions for the maintenance of equipment in the later stage.

## **References**

- [1] Li Wei (2018). Discussion on the application of smart power plant construction and intelligent power generation technology [J]. Automation application, 11, 82-86.
- [2] Tu Xuewei(2019), Zheng Yafeng. Discussion on the construction of smart power plants [J]. Automation Expo, 01, 29-31.
- [3] Yang Yan(2019). Key Technologies and Application Analysis of Integrated Big Data Platform for Smart Power Plants[J]. SME Management and Technology (Mid-term publication), 05, 130-131.
- [4] Zhang Fan(2017). Key Technology and Application Analysis of Integrated Big Data Platform of Smart Power Plant[J]. Huadian Technology, 39(02), 1-3+7+76.
- [5] Liu Hongtao(2018). Key Technology and Application Analysis of Integrated Big Data Platform for Smart Power Plants[J]. Internal Combustion Engines and Parts, 10, 220-221.
- [6] Shu Changqing(2018). Key Technologies and Applications of Integrated Big Data Platform for Smart Power Plants[J]. Telecommunication Power Technology, 35(07), 125-126.