Research on Information Management Model of Petroleum Enterprises under the Big Data Background

Jinquan Yang
Department of Computer Science, University of Xi’an Shiyou, Xi’an 710065, China
jqyang@xsyu.edu.cn

Abstract. Application of big data is analysed via the present information management of petroleum enterprises in this paper. The logical structure and application architecture are proposed based on hodiernal production, administration and management of oil firms in China and mainstream big data technology. Meanwhile, future applications of the information management are looked ahead under big data environment.

1. Introduction
Along with the widespread use of information technology, human society has entered a completely new time of big data, which brings more development opportunities to modern enterprises, but also imperceptibly makes them be challenged more. To achieve the goal of sustainable development against the background of big data, modern enterprises must make corresponding adjustments to the original information management model.

The application of big data analysis in the petroleum industry is in the experimental stage at present, with the researchers from companies, universities, and oil corporations. The internet construction of oil and gas production in Chinese petroleum enterprises, for which proposing the concept of "digital oilfield" will provide unified standards for the information system construction, and for which interfaces and the other relevant technical details furnish certain theoretical guidance, is increased and will have the material basis of big data analysis when the construction takes shape.

The development of big data receives wide consideration with an irresistible force all over the world. How to use and analyze the huge original data effectively, transform into the exploitable knowledge and value, and solve the problems in daily life and work becomes people's focus. The new information management pattern of oil enterprises based on big data, which will determine the success or failure of the application, has relatively changed accordingly.

2. Big data concept, characteristic, and purpose
‘Big data’, the definition that is given by McKinsey global institute, refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyse[1]. It has four characteristics: massive data scale, rapid data circulation, diverse data types and low value density, namely, the 4V characteristics of big data: Volume, Velocity, Variety and Value.

2.1. The usage of big data concept
Big data can be divided into such fields as big data technology, big data engineering, big data science, and big data application. At present, big data technology and big data application are considered more, meanwhile, the problems involving engineering and science have not been taken seriously. Big data
engineering refers to the system engineering of planning, construction, the operation and management of big data, and big data science focuses on laws of discovery and verification of big data and on relationship between the laws with nature and social activities during the development and operation of big data networks.

Internet of things, cloud computing, mobile Internet, Internet of vehicles, mobile phones, tablets, PC and all kinds of sensors around the world are sources and carrying ways of data. These data apply to sensor networks, social networks, Internet texts and documents, Internet search indexes, astronomy, atmospheric sciences, genomics, biogeochemistry, other complex and interdisciplinary scientific research, military reconnaissance, medical records, photographic video archives, and large scale e-commerce, and so on.

2.2. The impact of big data
For general enterprises, big data mainly reflects in analysis and use, and secondary development projects. According to the analysis of enterprise information, not only the hidden data can be mined, but customer source will be improved through these latent information, and entity selling. Secondary development of data has been extensive applied to network services. From summary and analysis of the information, personalized solutions conforming to the needs of customers, and a new way of advertising are shaped. Combine the products with services in accordance with the big data analysis is not accidental, which is always achieved by the leader of data era.

Above all, the application of big data not marks the progress of era, but encourages people to explore the deeper field. Besides, three characteristics of big data -- large scale, fast operation and data diversity -- also need understanding. On the basis of these aspects, the essence of data is easier observed, and software application platform is effectively operated in the meantime.

3. Information management status quo of petroleum enterprises
Problems existed in the information management at present are analysed in the following section.

3.1. Tardy archiving and poor integrity of enterprise data
As the working environment of petroleum enterprises is relatively dispersive, the management generally divided into headquarters, project department, and operation site. Therefore, the collection of data information is slow, and the data cannot be timely filed, resulting in poor data integrity. Petroleum enterprises are mainly engaged in exploration, petrochemical, sales, storage and the other business. Each link generates data volumes, leading to its dense and complex. With the advent of Internet and big data era, the types and quantity of data are also increasing. These will all hinder the collection, causing the poor data information integrity.

3.2. Low quality of data information processing and low utilization
It is important to provide data processing work as serving for the production management of enterprises. Oil enterprises have gradually established their own information system, however most focus on local business, without strengthening the cooperation with other departments, giving rise to low utilization, poor purpose and application performance and other problems of information processing achievements. It is negative to provide service for enterprises, impeding the development of data information management[2].

3.3. Lack of modern specialist of data management
Petroleum enterprises are in the developing stage of modernization, and one of the most important tasks is to cultivate modern professors, the same with data information management. It requires managers who have modern knowledge and high literacy to use advanced technology and equipment, maintain and update as well. Nevertheless, most of them have not majored in that. Lacking professional knowledge leads to low efficiency of the whole data management.
4. Information management model of petroleum enterprises under the big data background

4.1. Big data involved in petroleum enterprises

It is meaningful to apply big data for analyzing structured, semi-structured and unstructured data in upstream and midstream of oil and gas field production operation. Data diversity is significant in forecasting needs of business management. The potential application of big data in exploration and production is to discover quickly, reduce costs, improve drill safety and increase production.

Based on the information system of relevant petroleum enterprises developed early and the big data investigation on the domestic oil and gas exploration, development and production process in recent years, the data information should be analyzed from the following four aspects.

4.1.1. Big data analysis in oil and gas field exploration

Big data of oil and gas field exploration is an important content of cloud infrastructure of that. Considering about data management, storage condition and application field at home, for the characteristic of various data management department, scattered storage and diverse application needs, build big data application system of oil and gas field exploration in the country, and realize management of overall process in data acquisition - storage - pretreatment - sharing - applying, integrate with the data resources of oil and gas field exploration, mine valuable information of that, and provide a wide range of services. Big data of oil and gas field exploration includes seismic, heavy magnetic, geological and well logging data. Analyze and study the data structure of all kinds in oil and gas field exploration, especially semi-structured and unstructured data.

4.1.2. Big data analysis in process of drilling production of oil and gas field

Besides monitoring and forewarning in the process of drilling production, big data analysis can employ drilling one to execute design and simulation of drilling scheme, real-time monitoring and optimized decision of drilling engineering, in order to improve the success rate and efficiency.

For the characteristics of dispersed block, complex formation environment, scattered data storage and multiple application demand in the country, build the big data analysis in process of drilling production of oil and gas field, and realize management of overall process in data collection - storage - pretreatment - sharing - applying of that, integrate with the data resources of that, mine valuable information of that, and establish case base and knowledge base of that. Big data of drilling production of oil and gas field includes drilling, geology, logging, logging, and fracturing data. Analyze and study the data structure of all kinds in process of drilling production, especially semi-structured and unstructured data.

4.1.3. Big data analysis during development of oil and gas field

Big data analysis can help evaluate production processes of oil and gas companies. These analyses involve geospatial, pushing, and reporting information, which can make the enterprise exploit the oil and gas wells more intelligently in more competitive fields.

Enhanced oil recovery is the goal of many companies of oil and gas production. Big data can use seismic, drilling and production data simultaneously to provide reservoir analysis engineers with variations of reservoir timely, and production operator with reconstruction schemes of lifting methods. Big data can also be used to guide fracturing of shale gas.

Big data in the development of oil and gas field contains daily production, various laboratory analysis, geological, and fracturing data. Analyze and study the data structure of all kinds during development of oil and gas field, especially semi-structured and unstructured data[3].

4.1.4. Big data analysis in maintaining the operation of oil and gas field

Predictive maintenance is not a new concept for oil and gas companies, however it has not received deserved attention and well application. In upstream production process, if collect and analyze pressure, volume, and temperature together, and compare with damage history data of previous
equipment, the prediction can be automated, the same with midstream pipeline, especially for failure analysis of health, safety and the environment. Analyze and study the data structure of all kinds in maintaining the operation, especially semi-structured and unstructured.

4.2. Unstructured data analysis of petroleum enterprises
In the process of exploration, development, and production of oil and gas field will produce a large number of unstructured data, such as document, graphics, video, audio data. Most of these data are saved as material in previous petroleum enterprise, while part realize electronic, but are limited with store and query. The part is the valuable information resource of petroleum enterprises for studying the unstructured data and its statistical analysis model, determining data classification, analyzing the data structure, and providing strong support for data storage and use.

5. Information management framework of oil enterprise under the background of big data

5.1. Information management model under the background of big data
Traditional information management mode only suits for partial management and application of structured data, rather than a large amount of semi-structured and unstructured data. Based on the application and data analysis of petroleum enterprises in the early stage, the information management model of petroleum enterprises under the background of big data is analyzed.

5.1.1. Data level
Big data application is to obtain valuable information, especially prediction one, through high performance computing environment with data analysis and mining methods for distributed storage of multi-source massive data (including structured, semi-structured and unstructured), based on distributed storage management technology, such as Google's GFS and open source Hadoop which achieve the application of large, distributed, and vast data accessed; MapReduce and Spark programming model that realize data processing on huge computer clusters with data analysis and mining; BigTable which is implemented the storage management of the unstructured data to deal with massive non-relational data.

For structured data, each cluster node is stored on the basis of a unified data standard, such as drilling, basin, production, maintenance data. For unstructured data, such as graphics, images, video, audio, seismic exploration, electromagnetic exploration, and test and analysis results, Hbase, Mongodb and other unstructured database management are adopted[4].

The basic nodes of big data management are clusters or cloud computing services. To build big data for oil enterprises needs to construct their own data clusters of enterprises in different departments, or to make use of existed computing clusters. Each unit includes server of database, file, node cluster map, Web and the directory, responsible for data management of structured, semi-structured, and unstructured, for data access and load balancing between each system and different nodes with the coordinator. Each cluster connects through their gateway to the Internet.

5.1.2. Application level
Big data of oil companies is a vital content of cloud construction of the enterprise informatization. Considering about data management, storage condition and application field at home, for the characteristic of various data management department, scattered storage and diverse application needs, build big data application system of enterprise companies in the country, and realize management of overall process in data acquisition - storage - pretreatment - sharing - applying, integrate with the data resources of enterprise companies, mine valuable information of that, and provide a wide range of services. Each application subsystem of management is integrated on the PaaS cloud platform, which realizes service and function sharing on the basis of data sharing, and thoroughly solves the relatively isolated applications and data of subsystem.
5.2. Big data logical framework of information management of petroleum enterprises

On the basis of the deployment of big data cluster nodes for information management of petroleum enterprises, the data architecture for that is constructed basing on Hadoop and other mainstream big data analysis technologies, which consists five layers, namely, hardware layer, technology layer, data layer, mining layer and application layer (as shown in Fig.1)[5].

![Diagram of Big data logical framework of information management of petroleum enterprises](image)

Fig.1 Big data logical framework of information management of petroleum enterprises

The hardware layer basically includes the hardware facilities supporting the cluster node construction, containing various servers, storage devices, switches, routers. The technology layer includes support technologies of big data application, such as Hadoop, NoSQL, SqlServer, Oracle, virtual computing.

The data layer mainly includes the storage of structured and unstructured data, containing the data of exploration, drilling and completion production, development, production and maintenance of oil and gas field, and enterprise operation and management. For structured data, each cluster node adopts parallel database for management access, and unstructured data, Hbase or Mongodb storage management are adopted.

The mining layer provide the required data and services for the upper application system with the employment of data extraction and cleaning technologies of traditional data warehouse for structured data, and services for the upper application by adopting MapReduce, Pig, Hive, Mahout and other technologies to realize different types and depths of data mining for unstructured data.

The application layer mainly provides data support and services for the production management of petroleum enterprises (the production exploration, drilling, development, and maintenance of oil and gas, and enterprise operation management), management services, decision support and enterprise scientific research.

6. Big data application architecture of information management of petroleum enterprises

The nature of big data application architecture of information management of oil enterprise is to fully receive production data in different departments and operation of oil and gas, and to realize production data sharing and collaboration in different unit structured and unstructured oil enterprise, adopting mainstream technologies such as data storage and analysis without changing the present situation of the unit under the condition of existing database system. Essentially it has the function of mining valuable information from the production management data of massive petroleum enterprises and realizing the prediction. Meanwhile, it needs to realize retrieval, extraction, cleaning, and
preprocessing of data in traditional database, and to provide data support for different business subsystems. The application architecture is shown in Fig.2.

Fig.2 Big data application architecture of information management of petroleum enterprises

Production management systems of petroleum enterprises in Fig.2 are supported by the data layer and mining layer in Fig.1, including exploration production management, drilling production management, and the development management of oil and gas field, enterprise operation management, and subsystem of management service.

The application system architecture is above data layer and mining layer of big data application architecture of information management of petroleum enterprises. Data and information of the application system are from the oil and gas production, operation and management of data cluster nodes of various units and departments, in the meantime, data from system simulation analysis results give feedbacks to each database and the nodes, which further improve the amount of data and the information of data node, jointly advance the database of each node and data resources construction.

7. Conclusion

The core issue of big data is not data or technology, but people. This paper presents the logical framework and the application architecture of big data construction of information management in petroleum enterprises, and discusses the application and service mode of that under the background of big data. Nevertheless, the main difficulty lies in people, especially their views. Data needs sharing, which requires persons to change the perspectives. If the data that companies spend a huge sum of money for is not shared in time, it may probably sleep forever in the corner of a unit or a department, without providing valuable information for exploration of petroleum enterprises, production of oil and gas field, and enterprise management. It is actually a huge waste of investment funds for the country and the enterprises.

References
[1] James Manyika. The next frontier for innovation, competition, and productivity. 2011.5.
[2] Bo Wei. Research on cloud computing platform of big data of oil enterprises. Journal of Communication World, 2017.11(in Chinese).
[3] Daoyuan Lin. From enterprise architecture to the theory and practice of smart fields. Petroleum Industry Press, ISBN 978-7-5183-2063-9(in Chinese).
[4] Zhenji Gao. Research on big data architecture and application of geological survey of oil and gas resources. Geographic information world,2016.2(in Chinese).
[5] Zongze Jin. Research on big data analysis model of oil and gas exploration and development [D]. Northeast petroleum university,2015(in Chinese).