ISSN: 2664-9640

Integration of Big Data and BIM Based Construction Industry

Shikang Chen[1], Ziqi Zhou[2, *]

[1] School of Software Engineering, Xin Jiang University, Xin Jiang, China;

[2] School of Architecture and Engineering, Xin Jiang University, Xin Jiang, China.

* Corresponding Author

Abstract

Since the reform and opening up, especially in the 21st century, China is developing rapidly, and the accompanying construction industry has also experienced unprecedented prosperity. Over the past four decades, China has built 65 billion square meters of housing, especially in the core urban areas. This has also led to a series of problems: urban traffic congestion, too small building spacing affecting daylighting, inaccurate indoor positioning, etc. People have taken many methods to solve relevant problems, but they have neglected the key technology - big data, which has grown up with the construction industry. Big data collects data in billions. If it can be used in the construction industry, it will solve many thorny problems: solving the heat island effect in dense areas, analyzing where new buildings can be built to make better use of land resources, providing more accurate indoor positioning, etc. BIM, as a long used technology in the construction industry, is no longer unfamiliar to people. People regard BIM as a "database" in the construction industry. It can analyze all aspects of architecture from many different angles, but the application of Bim in architecture alone may not solve the long-standing problems faced by China's construction industry. After a round of innovation, it is necessary to carry out a new round of innovation in the traditional construction industry. That is to integrate big data in a real sense.

Keywords

Big Data,Architectural Engineering,Integrated Development,BIM.

1. Current Situation of Big Data and Construction Industry at China and Other Countries

1.1. Situation of Big Data Research

Since the reform and opening up, especially after entering the 21st century, the world has had a new round of scientific and technological revolution. It is different from the previous revolution. Its revolutionary period has developed rapidly. As a result, the world's global information has exploded. The "big" in big data is the most important symbol. It can no longer be measured by G or T. with the advent of the cloud era, the growth of big data has laid a solid foundation. The United States is the first country in the world to study and promote it as a national strategy. In 2011, South Korea put forward the plan of "smart Seoul 2015", which also promoted big data as a national strategy. With countries all over the world turning big data into a national strategy and big data gradually penetrating into the society, China has also started to march into the field of big data.

In 2014, the word "big data" appeared for the first time in the government work report, which comprehensively elaborated China's strategic layout for the direction of big data in the future. Because it started late, it should step up to catch up with the world trend and strive to become one of the world's leading players as soon as possible.

ISSN: 2664-9640

1.2. Situation of BIM Based Construction Industry

BIM was first proposed and put into practice in the United States. After a series of complex evolution, it was basically finalized as the BIM we are now familiar with in 2002. Since 2006, a series of BIM pilot projects have been deployed in Hong Kong, which will be regarded as the future construction industry norms. BIM in China is still a late start. In 2001, the government issued a document to support vigorously promoting the application of BIM Technology.

Although big data and BIM started relatively late in Mizoguchi, after more than a decade of development, China has also entered the forefront of the world in related fields. Although they seem to be unrelated, they have close internal relations. Therefore, if big data is combined with BIM, it will certainly shine brightly in the construction industry in the future, help to promote the development of the national construction industry, and provide a model and guiding significance for the integration of big data and other industries.

2. Problems in Chinese Construction Industry

Problems in Traditional Construction Industry The section headings are in boldface capital and lowercase letters. Second level headings are typed as part of the succeeding paragraph (like the subsection heading of this paragraph). All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. When receiving the paper, we assume that the corresponding authors grant us the copyright to use.

2.1. Problems in Traditional Construction Industry

The main problems existing in the traditional construction industry are as follows: the institutional obstacles in the construction industry have not been completely eliminated, and the transformation of the enterprise management mechanism has not kept pace with the development requirements of the market economy; The business philosophy, technical reserve, management level and financing capacity of the enterprise are still relatively backward. The extensive growth of the industry is largely maintained by the labor input and resource consumption. It follows a development path of high input, low output and low efficiency. [1] This exposes the construction crisis faced by the traditional construction industry in the past. The problems shown by this crisis are no longer similar to those shown by the software crisis.

The software crisis is due to the fact that the backward software development technology has been unable to meet the rapidly growing demand for computer software, in which the demand is the first, and the demand has forced the transformation of software development, so the discipline of software engineering was born. The vigorous development of the construction industry has created the construction demand, and a series of problems caused by the construction demand have contributed to the "construction crisis". It covers a wide range of aspects, involving production safety, environmental protection, sustainable development and other factors.

2.2. Problems in Construction Industry Based on BIM

BIM refers to building information model, which is a new tool for architecture, engineering and civil engineering. BIM has always been the source code of big data in the construction industry, because the work of each component in a construction project can be defined through BIM, and big data can be formed on the basis of BIM. Big data exchange can be more conducive to the construction enterprises to strengthen the construction quality management of construction projects. [2] Simply put, BIM is a "database" of the building itself, which contains all data

ISSN: 2664-9640

sources for the construction of the building, including construction cost, area and volume, impact effect and other data.

It is different from the previous construction industry. With the help of data, it will be more accurate at all stages and can predict the future. The core idea of BIM also lies in the design and construction of auxiliary building projects in the whole life cycle, integrating all the information required by the project to assist decision-making and more efficient management. [3]This information is reusable and maintainable. It is more like the "big data" of the building itself than the real big data. What it contains is limited to itself, so it can also be regarded as "small data".

3. Integration Analysis of Big Data and Construction Industry

The spring breeze of reform and opening-up has been blowing all over the country, and China's modernization has been put on the agenda, which is the direct driving force for the vigorous development of the construction industry and big data. Over the past decade, China's construction enterprises have been on the rise, and the growth rate has not decreased. Ministry of Housing and Urban-Rural Development, together with National Development and Reform Commission and 13 other departments, issued 《Guiding Opinions on Promoting the Coordinated Development of Intelligent Building and Building Industrialization》. It proposes to accelerate the coordinated development of new generation information technology and building industrialization technology, and increase the integration and innovative application of new technologies such as BIM and big data in the whole process of construction. [4]

Among them, big data is a data set whose scale is so large that it greatly exceeds the capability of traditional database software tools in terms of acquisition, storage, management and analysis. It has four characteristics: massive data scale, rapid data flow, diverse data types and low value density. [5] Many data will be generated in people's daily life, such as flight information, call records, transaction flow and so on. Whoever masters big data will take the initiative in this era, so big data is becoming more and more important.

3.1. Adequacy of Integration

The gross output value of China's construction industry has been growing continuously in the past decade (up to 2021), as shown in Figure 1.

Figure 1: Total output value of China's construction industry

According to the data, the total output value has increased from 1321.79 billion yuan in 2012 to 29307.9 billion yuan in 2021, and the year-on-year increase range can be obtained, as shown in Figure 2.

ISSN: 2664-9640

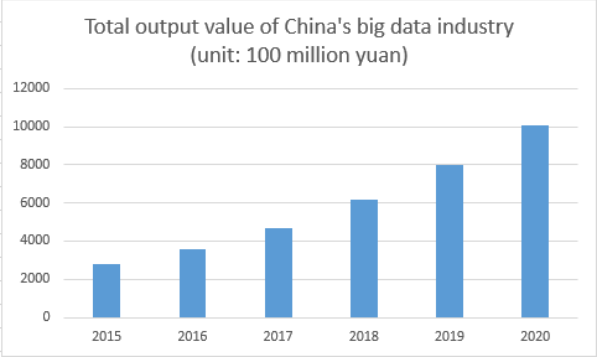

Figure 2: Year-on-year growth of total output value of China's construction industry

From 2012 to 2015, the situation fluctuated upward and downward, but never fell below 0. From the perspective of the huge construction industry, BIM can timely find data with important value to form relevant index data. Therefore, most construction units will conduct detailed management on the quality of the whole construction through BIM Technology. [6] As a model of the construction industry, BIM is regarded as a processing method of big data, which lays a foundation for the integration of big data and the construction industry, making the integrated two highly feasible.

Combined with the total output value of China's big data industry in the past six years (up to 2020), see Figure 3.

Figure 3: Total output value of China's big data industry

Big data will reach a total output value of RMB 1010billion in 2020, which is 3.83% of the total output value of the construction industry of RMB 2639274 billion. In addition, the growth rate of the gross output value of the big data industry will reach 26.25% in 2020, which is much higher than the 6.2% growth rate of the gross output value of the construction industry.

Therefore, the integration of the construction industry and big data has a high degree of sufficiency and predictability. On the one hand, the construction industry, which has existed for hundreds of years, needs a new power source in the 21st century to revitalize it. On the other hand, as a new industry, big data, it is actively integrating with various traditional industries, and it will radiate unprecedented vitality. Therefore, with BIM as the basic premise, the fusion will not be so strange to big data. It also has more sufficiency for practitioners in the construction industry.

3.2. Necessity of Integration

Under the situation that the country vigorously promotes the integration of big data technology and the construction industry, the integration of the two is mandatory and necessary. At present, traditional construction alone can no longer meet the growing construction industry. Traditional construction needs a long period from exploration and site selection to feasibility demonstration and then to surrounding planning. BIM can only help alleviate the problem to a

ISSN: 2664-9640

certain extent. Its role as a "database" does not lie in the scope it affects, but in itself. The real database can solve the above problems within a few hours. The site selection may only take a few minutes, and the feasibility demonstration is more reliable because of the support of big data. Therefore, the lack of database support is absolutely impossible and necessary.

On the one hand, BIM is only a small "database", so it has its own limitations and can not ensure that everything can be covered in the whole construction cycle. It only focuses on its own characteristics and ignores its impact on other aspects. This is like making a car behind closed doors, only caring about themselves but not the changes of external things, which leads to the unsustainable development of the construction industry. Big data is an extension of BIM, strengthening the scope and capability of BIM.

4. Technical System of Big Data and Selection of Data Storage

The era of big data is another brand-new revolution. It has subverted the traditional whole process of data collection and analysis. Therefore, computers are divided into computing intensive, I/O intensive and data intensive. Among them, data intensive computers mainly challenge the massive amount of data, the complexity of data and the rapid change of data. The technical system of big data may not overturn the entire construction industry at once, because the construction industry has already had BIM's "small database" as a transition, but if the database technology is fully integrated, it will also play a huge role in promoting and transforming the development of the entire industry. In the 1990s, with the birth of BI system, it has been able to help people with simple data management, but it is only a simple "database", because it can only add, delete, modify and query. But this has been able to solve a series of problems at that time.

4.1. Technical System of Big Data

The emergence of big data has subverted a series of traditional data processing technologies. For example, the change of big data acquisition mode has led to the rapid expansion of data scale. Compared with the traditional database system, its index, query and storage are facing severe tests, and how to quickly complete the analysis of big data is also a problem that the traditional data analysis methods cannot solve. [7] A series of technical systems of big data involve data collection, data storage, big data analysis and mining, and big data visualization, as shown in Figure 4.

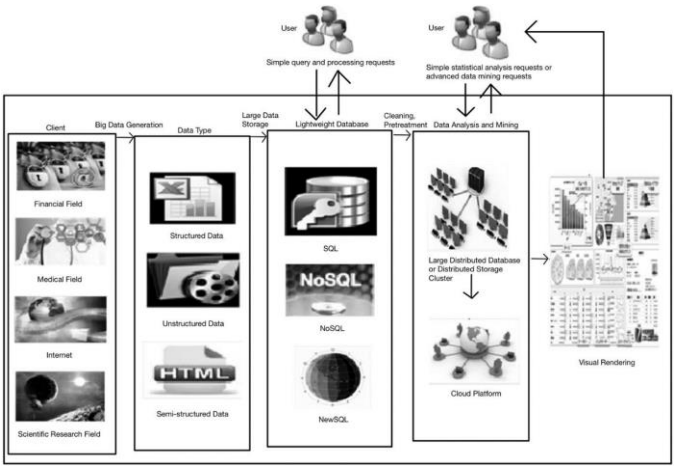

Figure 4:Big data processing technology system[7]

It can be seen from the figure that big data is generated by the financial field, the medical field, including the construction industry. The origin of these big data is due to the existence of various types of interactions and the spontaneous movement of people or things. They may be regular structured, irregular and unstructured, or the same interaction may generate both

structured and unstructured data, depending on the specific interaction subject. After the big data is generated, the big data operators will select the corresponding data types, which are structured data, unstructured data or semi-structured data. Then save the selected data type into the corresponding database. Finally, the corresponding database type is selected for data analysis and mining, and the corresponding algorithm is selected.

4.2. Source and Collection of Big Data

For the collection of big data, the construction industry has roughly had its own set of systems, which use BIM to collect and extract data. However, there are many types of data, such as structured data, semi-structured data and unstructured data. There is no real conflict between structured data and unstructured data. For the construction industry, both can be applied. Structured data includes high-speed storage application requirements, data backup requirements, data sharing requirements and data disaster recovery requirements. For the increasingly complex construction industry, many parts are involved, and the data access of each part must be completed within the specified time.

4.2.1. Qualitative Acquisition of Data

Qualitative data is obtained through non quantification. It does not depend on specific digital scales, but on people's subjective feelings. For the construction industry, quantitative data accounts for the vast majority, but qualitative data can not be ignored. BIM contains some qualitative data. Qualitative research is also called qualitative research, which can determine the essential attributes of things. At the beginning of the construction of the project, the common objective attributes of the target project and other projects can be extracted according to the shared BIM, and qualitative research can be carried out. Qualitative data is very easy to obtain. It is often descriptive, such as the architectural engineer's grand idea of the building, the laws and regulations that the building should comply with, the ecological environment around the building and the general description of Surveying and mapping. Some of the data can be directly obtained from BIM, but other data should be obtained through practice.

4.2.2. Quantitative Acquisition of Data

Quantitative data occupy almost the majority in the whole construction industry, because the precision and accuracy of the construction industry make quantitative data particularly important. If qualitative data is wrong, there is still room for maneuver, but if quantitative data is wrong, it will be a disaster. It not only has a great impact on the whole life cycle of the building, but also is very dangerous for people who use the building later. The acquisition of quantitative data comes from the samples of buildings. The specific data values are obtained through the accurate measurement of the samples. It can also be the data obtained by calculating the average value, variance and standard deviation of the samples. These data can be found in BIM, but for the data of the surrounding environment, it depends on other technical means.

4.2.3. Remote Sensing

Quantitative and qualitative data can also be obtained by Remote Sensing. Remote Sensing is a non-contact, long-range detection technology. Remote Sensing has spectral, spatial, and temporal characteristics. From the point of view of space-time characteristics, it conforms well to the space-time characteristics of the three-dimensional sketch of the building and is very important for later storage. Spectrum refers to the difference in brightness, hue, or cell values between different bands in a Remote Sensing image. [8] Different classes of objects can be distinguished based on different spectral characteristics, because each class has different spectral characteristics. Remote Sensing, therefore, has optical properties as a means of obtaining data. Secondly, its spatial characteristics are particularly important because all objects in real life are three-dimensional except for time, which is distinct from planar features. Planar studies only edge length and area, and the transition from two-dimensional to three-

dimensional is not only a traditional addition of one data type, but also a variety of additions. However, for the measurement of remote sensing spatial characteristics of buildings, shadows, patterns and other aspects need to be considered. Everything is moving, static is relative, and movement is absolute, which is also reflected in the time characteristics of Remote Sensing. The data measured by Remote Sensing is different for different time periods, so to get multiple sets of data for the same object, either compress the time or sacrifice the time. Compression time refers to the measurement of multiple sets of data in as short a time as possible, and the sacrifice time refers to the measurement again at the same time the next day and in the same surroundings. The former is often easier to achieve than the latter.

4.3. Analysis and Mining of Big Data

Big data analysis and mining is the biggest difference for BIM. Although BIM itself can also perform relevant analysis and visualization, its ability is far inferior to that of big data analysis and mining. On the one hand, it only uses the data of the building itself to analyze the building itself, which has certain limitations. It can not well analyze the impact of the surrounding environment on the building and the impact of the building itself on the surrounding environment. At present, when sustainable development is required, we should also pay attention to the integration with the surrounding environment, which is also the pursuit of architectural designers themselves.

Big data analysis and mining in the construction industry first extracts data from the database, then performs corresponding transformation, and performs data cleaning and loading. Then, the preliminarily processed data is put into the data warehouse, and then the data from the data warehouse is further processed into the data mart. The data mart can analyze, mine and other process the data. Finally, the analyzed data can be used to obtain its behavior analysis The change trend and future forecast are further refined into decision-making application support and analysis result visualization according to the relationship between the three. Finally, the analysis result is visually output and presented to users.

4.3.1. Quantitative Acquisition of Data

Artificial neural network can simulate the neural network in human brain through a large number of simply connected artificial neurons. It has been widely used in pattern recognition, machine vision, associative memory and other fields. Of course, it can also be applied in the field of data mining. In the neural network model, the neuron first gives the weight value 'w_{nj}' to the input value 'x_n' of the external environment or other neurons, and then calculates it through the conversion function Z_j and the response function f (Z_j). Finally, the result Y_j is output to the external environment or other neurons. [9] The operation diagram is shown in Figure 5.

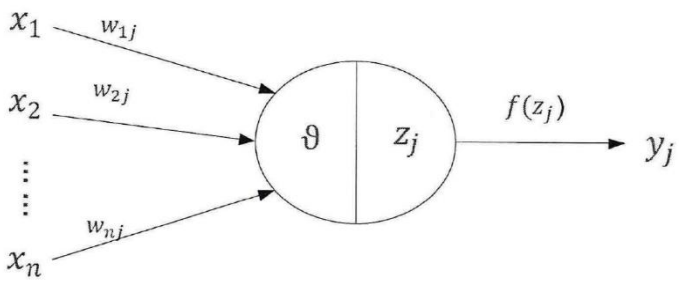

Figure 5:Structure of neurons[9]

It establishes the connection with big data through the input layer, and outputs through the interface connected by the output layer after a series of calculations in the hidden layer. Among them, BP neural network in artificial neural network has better performance, and has a BP theorem: given random value ε> 0, for random continuous function f, there is a three-layer

feedforward neural network, which can is infinitely close to the 'f' function at any square error precision of 'ε'. So it has better big data mining potential.

4.4. Visualization of Big Data

Visualization of big data is very simple. It relies on the big data mining and analysis in the previous step, because the data in the construction industry is massive, and it is meaningless to present it to the construction personnel or project managers. Just presenting the stored data intact is not big data, it is only the most original data access. The most important difference between big data and other data is that big data is intelligent. The intelligence mentioned here refers to its ability to analyze and mine data. Therefore, it is the most important and critical to present the analyzed and mined data to everyone in a visual way.

At present, there are four types of data visualization: multidimensional data visualization, time series data visualization, network data visualization and hierarchical information data visualization. Most of the data in the construction industry are 3D data. On the one hand, you can choose to display the 3D data directly, on the other hand, you can choose to reduce the dimension for visualization, which can better reflect the data required by the builders.

4.5. Selection of Data Storage

For the massive data, the key is to select the storage method. Both light database and large database storage platforms can be selected, but the latter should be selected for the construction industry. Although the former can also store structured data and query structured language, it can not query a large amount of data for SQL. For the query of a large amount of data, simple query needs to quickly return results, application of unstructured data, etc., so the relational database used for big data storage needs to make different improvements to meet the storage and query requirements of big data. [7]SQL can only return a small amount of data, but can not return a large amount of data. Therefore, it needs to be slightly modified before querying.

At present, typical big data storage platforms include Info-Brignt、Hadoop（Pig and Hive）、YunTable、HA-NA, exadata, etc. the above databases, except Hadoop, can meet the online analysis requests of big data. [7] Therefore, the database selection of construction industry can start from these databases.

5. Advantages and Disadvantages of Integration of Big Data and Construction Industry

There is no perfect thing in the world, and the integration of big data and the construction industry is not once and for all. It must also have many shortcomings in the initial stage. It is not the key to solve the shortcomings. The key lies in how to find them. Only in this way can they flourish. Keywords such as 'digital expression', 'resources'and 'data' link it with big data, opening up new ideas and methods that have not been tried before. The advantages and disadvantages of fusion still need to be determined according to the depth and breadth of fusion. Predictably, the contradiction will always exist between the two. In the early stage of integration, BIM will be excluded in some aspects due to its similarity with big data.

5.1. Advantages after Integration

The integration of the two on the one hand it can solve many problems about the construction industry, on the other hand, it can also improve the efficiency of construction. The construction cycle from the beginning to the end of the design and construction can be made faster and more efficient through the addition of big data, advocating the integration of the two can not only drive the development of big data industry, but also pull the development of the construction industry, with more advanced technology into it, there will be more big data practitioners into

the construction industry, and then along with the deep integration of the two to stimulate economic development. BIM also consumes less energy than the traditional construction industry. In the construction of resource-saving and environment-friendly society, big data is especially important. Under the slogan of conforming to nature, protecting nature and respecting nature, harmony between human and nature has become the main theme of this era. China is gradually eliminating high energy consumption and low capacity buildings on the one hand, and actively developing green buildings on the other, and the premise of green buildings cannot be achieved without the addition of big data.

5.2. Disadvantages after Integration

Management is one of the most important aspects, and the lack of knowledge and experience of the management will lead to a series of unpredictable and terrible consequences. In comparison with the software crisis, one of the most important aspects is the crisis of the management itself. First of all, it does not understand the main contradictions faced by the transformed industry, and still bases its solutions on the pre-transformation contradictions, which often leads to fatal mistakes. On the other hand, there are also a lot of people resisting big data, and a series of ethical problems caused by big data have emerged in recent years. The equality issues raised by big data applications mainly include the data divide and data hegemony. Equality is one of the most dominant values in human society today, and the concept of justice in modern society is based on equality. The data divide can lead to a deepening of the fragmentation between different regions and groups, highlighting the role of technology as a barrier to segregation and exacerbating social inequalities and divisions. [10] Big data can also be a double-edged sword to draw; on the one hand it can be a good catalyst for the development of the construction industry, and on the other hand it can be a source of hindrance to its development. The development of things is never smooth, it is always moving in the direction of the wheel of history, wave forward and spiral up.

6. Prospects for the Future Industry

In the era of widespread and continuous development of big data technology, new technologies and concepts represented by "Internet +" have opened up new horizons for the management of the building operation stage. [11] It will shine in the future development. The use of big data in the construction industry will make the whole society gain. It can analyze various favorable as well as unfavorable factors in a very short period of time. It can solve many of the development bottlenecks faced by cities today, and it can also better help the construction industry to conduct more in-depth research that will drive its development. It can also design and develop an energy consumption data management platform for the building operation and maintenance period on the basis of key technologies to provide a support platform and tools for the complete process of monitoring, managing, and controlling energy consumption management data during the building operation and maintenance period[12]. These are all attempts relying on big data, which are beneficial for the industry to develop relevant methods and technologies.

But on the other hand, opportunities and challenges coexist. The integration of big data with industries is not high enough. The corresponding image can be drawn according to the percentage of integration of each industry with big data, as shown in Figure 6.

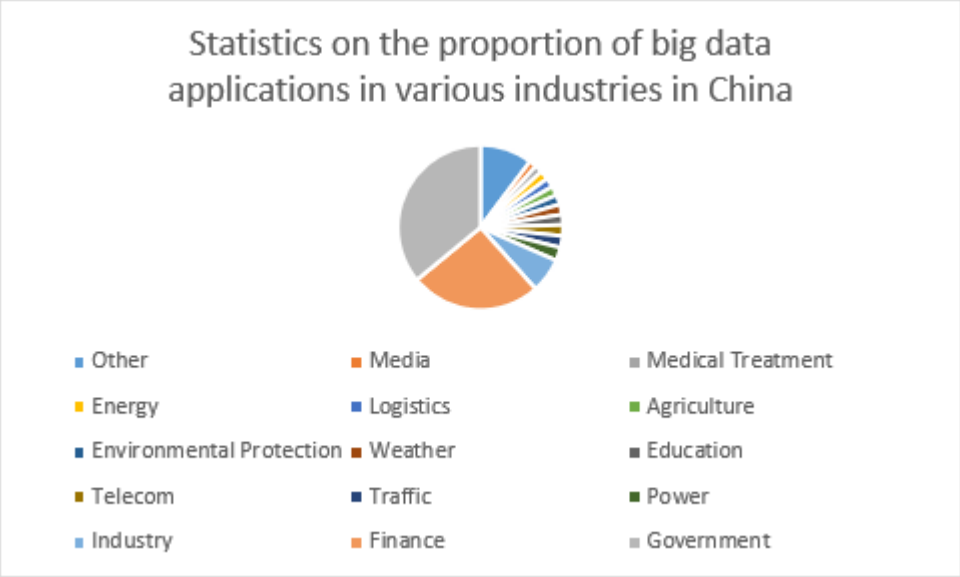

Figure 6:Statistics on the proportion of big data applications in various industries in China

It can be seen that big data has the closest relationship with the government, followed by the financial sector and then industry. There is no excuse for the close relationship between big data and the government, because it assumes the important responsibility of maintaining national security and social stability. The close relationship with finance and industry is also deserved, both are important areas for the survival of the country. But compared to other industries, the proportion of big data is still relatively low, especially in the construction industry, which is attributed to only a part of the other. However, we can conclude that although the current proportion is small, but the development potential is huge, so we still need to be confident in the construction industry.

To combine the construction industry with big data will be a very simple but very difficult thing. With the continuous development of urban construction, modern building operation and maintenance management has become a very complex management model, and there are different characteristics of buildings between cities in China, and their management regulations and standard models have their own characteristics. If we want to vigorously develop the integration of big data and BIM, it is inevitable to cultivate relevant talents, and the cultivation of talents requires continuous trial and error and spending a lot of time.

However, the integration of big data and the construction industry will be a trendy event, and big data will be more involved in the construction industry in the future.

References

[1] Commentator of this journal: Getting Out of the Traditional Construction Industry, Construction Enterprise Management, Vol. 24 (2008) No.5, p.9.

[2] Wei Gao: Application of BIM and Big Data in Quality Management of Construction Engineering,Jushe, Vol. 40 (2020) No.30, p.102-103.

[3] Li Shen: BIM-based Primary Research on the Big Data of the AEC Industry (MS., Southwest Jiaotong University, China 2016), p.1.

[4] Information on: https://baijiahao.baidu.com/s?id=1673461971185590075&wfr=spider&for=pc

[5] Yuanwang Xia: Big Data Thinking in the Big Data Era, Henan Daily, Vol. 69 (2017) No.5, p.9.

[6] Yong Wang: BIM and Big Data Technology Application in Construction Project Quality Management, Marketing Management Review, Vol. 29 (2021) No.6, p.115-116.

[7] Yu Peng , Jingyue Pang, Datong Liu, Xiyuan Peng: Big data: connotation, technical framework and its development, JOURNAL OF ELECTRONIC MEASUREMENT AND INSTRUMENTATION, Vol. 29 (2015) No.4, p.469-482.

[8] Changjun Li: Design, Research on Cultivated Land Extraction Technology of Agricultural Remote Sensing Images Based on Deep Learning (MS., University of Science and Technology of China, China 2018), p.1.

[9] Liyan Huang: Research on Passenger Flow Analysis of High Speed Railway Based on Big-Data (MS., Southwest Jiaotong University, China 2017), p.1.

[10] Yi Zhao: The Research on the Ethical Issues of Big Data in Application (Ph.D., Dalian University of Technology, China 2021), p.1.

[11] Jingwen Huang: Research on Building Operation and Maintenance Management used in BIM and Big Data (MS., Beijing University of Posts and Telecommunications, China 2017), p.1.

[12] Yunyi Zhang: BIM-Based Public Building Maintenance Energy Big Data Management and Analysis (Ph.D., Tsinghua University, China 2020), p.1.