Evolution of Business Intelligence: An Analysis from the Perspective of Social Network

Hanqing HU, Mengyue YIN, Jianling LI*

Abstract: Based on CiteSpace, Pajek and other software, this paper makes a visual analysis of the knowledge graph of the related literature of Business Intelligence and explores the future development trend of business intelligence. Taking the core periodicals of CNKI as the data source, key words are drawn and analyzed with the help of software. The total number of articles was 2938 from 2006 to 2020, and the number of articles published in the past 15 years was gradually levelled off. Among the 607 researchers, Yang Bingru is the representative; there are 424 journals, Journal of Information is the first, and 787 keywords are the most frequently used data mining. Our country still needs in-depth research in the field of business intelligence. Through the atlas, it directly shows that big data and machine learning are the frontier hot spots of future development, which provides research direction for researchers.

Keywords: Business Intelligence; Knowledge Atlas; Research Status; Social Network Analysis

1 INTRODUCTION

The advent of the information age has brought many new opportunities and challenges to the development of enterprises, and the strength of information processing and utilization capabilities has become the key to the success or failure of enterprises. Business Intelligence software is gradually derived with the strengthening of enterprises' demand for information utilization. According to the survey [1], enterprises spend about 150000 - 300000 dollars on business intelligence software every year, which only includes simple use of rights and maintenance, and other expenses are not included. Everson said: for every dollar you spend on business intelligence software, you have to pay five to seven times the relevant expenses. Despite the high cost, a large number of enterprises are willing to pay the money, which also reflects the actual value of business intelligence in the application process.

The concept of Business Intelligence was not proposed until recently, it was first proposed by the Gartner Group in 1996 [4]. The Gartner Group defines Business Intelligence as the following: Business Intelligence describes a set of concepts and methods that assist business decision making through the application of fact-based support systems [5]. Chinese scholars have been actively studying Business Intelligence and have made some achievements in theoretical research and practical application. As can be seen from Fig. 1, the study of Business Intelligence in China has gradually levelled off, the trend line obtained by Excel analysis is $y = -62.84 \ln(x) + 312.75$, $R^2 = 0.9162$, which is more reliable than other trend lines. In 2006, the number of papers published in 2010 was more than 200, the same as in 2011; the number of papers published in 2020 fluctuated. According to the current situation of papers published by CNKI, the number of papers published in 2020 is predicted to rise, which is in line with its stable trend. This also shows that China's research on business intelligence is in the development stage. Based on the published literature as the research object, this paper systematically summarizes the research situation of Business Intelligence, in order to provide reference for researchers' future research.

2 LITERATURE REVIEW

Through sorting out the existing literature [6-8], it is found that there are only two types of representative views at present. One is the perspective of the business community represented by IBM. They think from the perspective of entrepreneurs. They believe that business intelligence is mainly based on the information collected by IT technology for the purpose of improving enterprise performance. Another academic point of view is represented by Osizak, who thinks from the perspective of scholars that business intelligence is a collection of all methods that can help enterprises make decisions and belongs to the tool for enterprise decision-making.

Based on the above two views, we can know the importance of data information for Business Intelligence. Data resources are the intangible asset of the enterprise, and Google is able to shine in that era because it has a lot of data. When the technological advancement level of each enterprise is basically the same, whoever has more data will have the advantage of breaking through first. Google saw this advantage and quickly broke through before other enterprises grasped a large amount of data, but from this other enterprises also realized the importance of data, so this advantage was quickly pulled closer in the second year. Since the emergence and development of Business Intelligence system, its system has gradually tended to mature and play a role in combination with various fields [9-12]. Therefore, based on the published literature as the research object, this paper systematically summarizes the research situation of Business Intelligence, in order to provide reference for researchers in the future.

![Annual publication trend chart](image)
3 RESEARCH METHOD

Social Network analysis is a method developed by sociologists based on mathematical methods, graph theory, and Quantitative analysis [13]. Originally conceived in 1930 by Emile Durkheim (1879), the method has developed rapidly in recent years and has found wide application in many fields. CNKI is the largest Chinese database in China, covering more than 90% of the core journals. In order to make the retrieved journals authoritative, the selected databases are "SCI", "EI" and "CSSCI". Because the core technologies of Business Intelligence are Data Warehouse, On-Line Analytical Processing, and Data Mining [11], in this paper, the subject of "Business Intelligence" OR "data warehouse" OR "on-line analytical processing" OR "data mining" is searched in the database. The searching period is from 2006 to 2020 (the searching time is 2021 May 30), all other retrieval conditions are system default. A total of 4778 papers in Chinese journals were retrieved, and 2938 papers were rejected.

By studying the network relationship, the macro large network is connected with the micro small points, so as to explore the deeper reasons that have not been consciously recognized [14]. In this paper, CiteSpace and pajek software are used to construct the network, and the medium neutrality of keyword network is used to analyze the relationship between each node in the network and the relationship between the node itself and all points in the network.

4 RESEARCH RESULT AND DISCUSSION

4.1 Author Collaboration Characteristics

With the help of a knowledge atlas of the citespace software (Fig. 2), it is possible to visualize the degree of collaboration and closeness among the various authors. It is calculated that the number of nodes is 607, the number of connections is 390, the network density is 0.0021, and the density value is low. This shows that the connection among the nodes in the network is not very close. At the same time, it can be seen that the co-operation among the authors in this field is loose and the overall structure is not tight, there are multiple teams working on the study, and authors in this field is loose and the overall structure is not tight. At the same time, it can be seen that the co-operation among the authors in this field is loose and the overall structure is not tight, there are multiple teams working on the study, and teamwork can range from two to three people or more, but there is no collaboration between the different teams. There is more collaboration between the teams such as Yang Bingru, Zhang Kejun, Song Zefeng, Qu Wenlong and Lu Jianfeng, Chang Hao, Yan Wei. At the same time, the scattered independent nodes in the graph indicate that some scholars are doing their research independently. The most prolific authors were Zhang Yufeng, 21 in total, 16 of them had more than 7 posts (Tab. 1). It can be seen that the most prolific authors were all team players. In addition, Li Yongnan's Burst intensity is 7.47 by using the Burst analysis function of citespace, and his articles are collected in 2017 and in 2020, followed by Yang Bingru with Burst intensity of 5.18 and time of publication in 2008 and in 2010.

4.2 Hot Periodicals Characteristics

Using pajek software to analyze journals, 2938 articles of this paper are published by 424 journals, 33 journals with more than 20 articles. Journal analysis provides the direction of knowledge accumulation for the research in this field; to some extent, it reflects the theoretical and practical value of the research field. Because of the length, this paper lists only the top 10 hot journals with the largest number of articles published. The Information Journal has 138 articles published, accounting for 4.70% of the total, the second were computer research and development (120, 4.08%) and computer engineering (117, 3.98%). The Journal of Computer Science with 6.477, the Journal of Software (4.068) and the Journal of Information Theory and practice (3.276) were the most influential factors. It shows that the research of Business Intelligence in China has certain authority.
4.3 Research Topic Characteristics

4.3.1 Keyword Ranking Analysis

After summarizing, replacing and integrating the keywords provided by the selected literature authors, the results show that the top 10 keywords are as shown in Tab. 3. The highest keyword frequency has "data mining"; it is 2048, and 256 words with word frequency ≥ 5 account for 44.36% of the total. You can see the three core Business Intelligence technologies: data mining, data warehouse and online analytical processing are all in the top 10.

<table>
<thead>
<tr>
<th>Item</th>
<th>Keyword</th>
<th>Frequency</th>
<th>Percentage / %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Data Mining</td>
<td>2048</td>
<td>15.7019</td>
</tr>
<tr>
<td>2</td>
<td>Association rule</td>
<td>214</td>
<td>1.6407</td>
</tr>
<tr>
<td>3</td>
<td>Data Warehouse</td>
<td>203</td>
<td>1.5564</td>
</tr>
<tr>
<td>4</td>
<td>Cluster</td>
<td>195</td>
<td>1.4951</td>
</tr>
<tr>
<td>5</td>
<td>Big data</td>
<td>159</td>
<td>1.2190</td>
</tr>
<tr>
<td>6</td>
<td>Decision-making tree</td>
<td>79</td>
<td>0.6057</td>
</tr>
<tr>
<td>7</td>
<td>Library</td>
<td>76</td>
<td>0.5827</td>
</tr>
<tr>
<td>8</td>
<td>OLAP</td>
<td>59</td>
<td>0.4523</td>
</tr>
<tr>
<td>9</td>
<td>Rough set</td>
<td>51</td>
<td>0.3910</td>
</tr>
<tr>
<td>10</td>
<td>Data flow</td>
<td>46</td>
<td>0.3527</td>
</tr>
</tbody>
</table>

4.3.2 Keyword Co-occurrence Analysis

Key words network is drawn according to co-occurrence Matrix (Fig. 3), and the size of nodes is set according to the value of Betweenness, which is an important index to measure the ability of nodes to communicate information in social network, reflect the node as a "bridge" to the importance of the degree. Betweenness is defined as:

\[ C_B(v) = \sum_{s,t \in V \setminus V(v)} \frac{\sigma_{st}(v)}{\sigma_{st}} \]  

where \( \sigma_{st} \) is the number of shortest path through node \( v \).

The most central keyword is "data mining" with a value of 0.35, and according to Fig. 3, the most interesting and influential hot topics in this field are mainly focused on Business Intelligence technology.

\[ d(v) = \frac{2M}{N(N-1)} \]  

\[ \bar{d} = \frac{1}{N} \sum_{v \in V} \text{deg}(v) \]  

where \( d(v) \) represents the average network density, \( m \) represents the number of edges in the network, and \( n \) represents the number of nodes in the network. Finally, the network density value is 0.192, which indicates that the nodes in this paper are closely related to each other. According to the Eq. (3), the average degree of nodes is 29.21, which shows that some nodes have many relationships and some nodes have very few relationships.

\[ C_{v_2} = \frac{n}{C_k^2} \frac{2n}{k(k-1)} \]  

The aggregation Coefficient (Eq. 4) represents the degree of node aggregation in a graph. The value obtained in this paper is 0.53, which represents the relative aggregation of nodes.
By sorting the keywords in the graph, we find that among the 76 keywords, 70 are related to technology, accounting for 92%, and 6 are related to application environment, accounting for 8%. China's Business Intelligence research focuses on technical issues. Typical keywords for both areas are shown in Tab. 4.

![Image of Keyword co-occurrence Network Diagram](image)

**Table 4 Keyword classification**

<table>
<thead>
<tr>
<th>Research Contents</th>
<th>Typical Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology</td>
<td>Data Mining: Association rule;</td>
</tr>
<tr>
<td></td>
<td>Data Warehouse: Cluster; Big data; Decision-making tree and so on</td>
</tr>
<tr>
<td>Application Environment</td>
<td>Library; Electronic Commerce; MicroBlog: Commercial Bank, Competitive Intelligence; Higher school library</td>
</tr>
</tbody>
</table>

4.4 The Hot Frontier of Business Intelligence

With the help of the keyword emergence function of CiteSpace, can be reflected the heat of a research direction in the field, that is, the research frontier or research hotspot (Fig. 4). In the early stage, the domestic research on Business Intelligence was in the exploratory stage, which is limited by professional knowledge. Researchers at home and abroad focused on the research of data warehouse and other technologies, and business intelligence technology was mainly used to deal with structured problems [15]. Since the rise of cloud computing and big data technology in 2012, Business Intelligence has started to play its role by combining with new technologies [16-18]. Since then, the combination of Business Intelligence and big data has kept the forefront of development. In 2017, Business Intelligence has been integrated with machine learning, and the research direction has gradually become mainstream, leading the development of the industry.


![Image of Keywords with the Strongest Citation Bursts](image)

4.4.1 Business Intelligence and Big Data

In the age of DT (Data Technology), decision-making using the analysis of Data volumes can increase the effectiveness of the results. In the process of analyzing data content, deep learning and knowledge computing are its foundation, visualization is its key technology; in addition, visualization is also the key technology of presenting data analysis results [19]. In the background of big data, Business Intelligence has multi-disciplinary cross-attributes, in the actual application of cross, complexity, novel, flexible and configurable [20-24].

Zelenka M [25] analyzes the effectiveness of business intelligence tools and claims that it depends on many factors. One factor of major importance is data quality. Ram Jiwat [26] believes that although big data analysis is very important, there is little research on the impact of using big data analysis for business intelligence purposes. This study fills this gap in knowledge by examining the role and implication of Big Data analytics on business intelligence for the data collected from Social media channels in China. Xiao Wenlong [15] and others use the CiteSpace software to analyze the development trends of Business Intelligence and big data at home and abroad, analyze the differences between domestic and foreign research, compare and analyze the insufficiency of our country in this research field, and conclude that foreign research focuses on management and decision-making, while domestic research focuses on technology and applications. Future research will focus on artificial intelligence and business analysis. Debotroli S [27] compared business intelligence with big data skills based on text mining research of recruitment advertisements. Scholars have fully studied the supporting role of big data for Business Intelligence decision-making. At the same time, the future development of big data is envisaged.
Based on the previous analysis, this paper concludes that the combination of big data and Business Intelligence has the following problems: (1) there are limitations when using, enterprises cannot give full play to the existing data. (2) Many enterprises do not have "data consciousness", do not know the initiative to collect data or cannot collect data. (3) There is no innovative thinking in the application process and no innovative application of the existing data. (4) The data cannot be updated on time, the database cannot be updated in real time, using some meaningless data. In order to solve the above-mentioned problems, it is necessary to strengthen the construction of Business Intelligence system and improve the overall level of its application [28, 29].

The era of big data brings both opportunities and great challenges. In order to give full play to the value of big data, we must solve the basic challenges. Only in this way can we bring breakthrough progress to the use of big data [30]. In addition, how to make the perfect combination of big data and enterprise operation and development is a problem that every enterprise needs to think about and constantly improve, and the construction of a perfect business intelligence system plays a supporting role [31-34].

4.4.2 Business Intelligence and Machine Learning

Machine learning is the core of artificial intelligence and the fundamental approach to make computers intelligent [35, 36]. The combination of Business Intelligence and machine learning is also the combination of artificial intelligence. This paper predicts that this will become a general trend of future development and bring more profound changes to Business Intelligence technology and enterprises. At present, machine learning is mostly based on labelled big data to train the model for machine learning, but some small enterprises may not have the strength to do a large number of data markers based on specific business needs. Therefore, reducing the dependence of machine learning on big data is the hot direction of current research. If this problem can be solved, the strong artificial era may come [37].

Jarrahi, MH [38] says that artificial intelligence has penetrated into many organizational processes, resulting in a growing fear that smart machines will soon replace many humans in decision making. To provide a more proactive and pragmatic perspective, he highlights the complementarity of humans and AI and examines how each can bring their own strength in organizational decision-making processes typically characterized by uncertainty, complexity, and equivocality. He Qing et Al [39] say that many traditional machine learning algorithms are no longer suitable for the current situation under the background of the explosive growth of data volume and the variety of data characteristics, they need to be improved in line with the changing background of the Times. In view of the large data background, Zhao Lin [40] discusses the application scheme and related technology of enhanced learning in business intelligence in machine learning method. By using reinforcement learning algorithm, aiming at the characteristics of business intelligence system in large data scenarios, this paper explores the construction method and application process of business intelligence system with reinforcement learning. Doel Shivendra [41] believes that machine learning is an applicative science that has optimized algorithms and models for data analysis. He focuses on machine learning and its assortment with Artificial Intelligence (AI) techniques for the purpose of data analysis. Based on the previous analysis, the following problems exist in the combination of machine learning and Business Intelligence: (1) Machine learning relies heavily on data, therefore, the characteristics of data itself make it difficult for traditional machine learning algorithms to process and analyze effectively. (2) Human-computer cooperation cannot achieve the expected effect, human beings cannot really let go of the machine, and the cooperation mode between the two cannot be determined. How to solve the above two problems needs more scholars to conduct in-depth research.

Although the theory of machine learning has been successful, there is still a large number of problems that have not been solved, which requires deeper mathematical knowledge and innovative thinking, and still needs continuous in-depth research [42]. At present, China has taken the development of artificial intelligence as a national strategy, and has issued a lot of policies to support the new technologies. Related basic and application researches are also emerging one after another. Business Intelligence will continue to improve with the common development of artificial intelligence.

5 CONCLUSION

Taking 2938 documents on Business Intelligence collected by CNKI from 2006 to 2020 as the research object, this paper reveals the research power and development trend of Business Intelligence in recent 15 years, and summarizes the basic survey of Business Intelligence research in China. Although the research on Business Intelligence in our country started late, it has its own architecture now and it tends to be stable, with an average annual volume of more than 100 articles. The author's research work is either carried out jointly with his own team or carried out alone. Third, the published journals of Chinese literature are authoritative, which also represents the international status of Chinese research; fourth, the research on Business Intelligence in China is mainly focused on the improvement of technology, and it is found that the focus of the following research will be on the combination of big data and machine learning.

According to the above analysis, the development of Business Intelligence is in a flourishing period at present. The author puts forward the following suggestions for the future development of Business Intelligence: (1) scholars should establish or strengthen cooperative links and conduct academic research together, share their experiences. (2) It is suggested to broaden the research field of vision and carry out innovative research on the development of bidding abroad. (3) It is suggested that the combination of Business Intelligence with big data and machine learning should be vigorously developed to promote technological change and assist enterprises in decision-making more efficiently. (4) It is suggested that the innovation of application scenario should be taken into account while paying attention to the development of
technology, so as to develop more possibilities for the application of Business Intelligence system.

Acknowledgements

(1) Beijing Municipal Social Science Foundation key project, 18YJAJ003.
(2) Top-notch plan project for strengthening schools with talents, BPHR2020BS01.
(3) A series of research projects of the service center of China Association for Science and Technology, 2020GGFWTLS005.
(4) Graduate Course construction project of Beijing Information Science & Technology University, 5027011033.
(5) Teaching Reform Project of Beijing Information Science & Technology University, 5112110857.

6 REFERENCES

[27] Deboldt, S., Muller, O., & vom Brocke, J. (2014). Comparing Business Intelligence and Big Data Skills A Text Mining Study Using Job Advertisements. *Business And


