



Trends and Evolution of Research Hotspots in Organizational Innovation Under the Background of Big Data—A Knowledge Graph Analysis Based on Citespace

Jia Liu(✉)

School of Business Administration, Guangdong University of Finance and Economics,
Guangzhou, China

18810627786@163.com

Abstract. In the era of big data, enterprise competition is complex and diverse. Innovation is an important skill for organizations to survive, but the relevant research in academia is not mature. There are problems such as scattered topics and traditional research ideas, etc. This paper uses citespace bibliometric tools to visually analyze the existing important literature such as author cooperation network and keyword co-occurrence, aiming at discover research trends in this field and provide direction for future research. The research findings are as follows. First, the research of organizational innovation in the context of big data can be divided into three stages, 2000–2011 as the starting stage, 2012–2014 as the slow development stage, and 2014–2021 as the rapid development stage. Second, the research hotspots in the field of organizational innovation under the background of big data from 2000 to 2021 are about the organic combination of information technology, knowledge management, enterprise management, etc. The most important foothold in the research field is organizational performance, including innovation performance, management performance, financial performance, etc. Third, exploring the influencing factors of organizational innovation in the Chinese context and the dynamics of organizational innovation will be the future research direction.

Keywords: Big Data · Organizational Innovation · Citespace

1 Introduction

In the information age, the amount of digitized information is rapidly increasing, and the society has accelerated into the era of big data. Big data is a collection of data, which has a large volume, a fast update speed, and contains a variety of specific points. It is easy to organize and collect, and can be produced at low cost. High value, social and economic development presents a new trend of intelligence, digitization, and networking. In the data era, multiple cooperative and competitive relationships have been formed between enterprises and enterprises, between enterprises and organizations, and between organizations [4]. The complex and changeable economic forms in the era of big data promote

organizational innovation to improve organizational efficiency. Accordingly, organizational innovation in the context of big data has attracted more and more attention from the academic community. Organizational innovation was proposed by Schumpeter. It has gone through the initial development stage, growth evolution stage, mature development stage and new transformation stage. With the further development of the era of big data, organizational innovation has been transformed into digital technology. However, related research is not mature, and there are problems such as scattered research topics and traditional research ideas. This paper uses CiteSpace to carry out statistical and quantitative analysis on the literature from 2000 to 2021, and uses word frequency analysis, co-occurrence analysis, cluster analysis, emergent analysis and other methods to study the current situation and development trend of organizational innovation research under the background of big data. Provide direction for future research in the field.

2 Research Method and Data Source

CiteSpace, also known as citation space, was developed by Professor Chaomei Chen from the School of Information Science and Technology, Drexel University (Philadelphia, PA, USA), and grew up in the context of metrology and visualization to discover hidden knowledge. Bibliometric visualization tool. Its visual structure shows obvious scientific distribution and laws, so the result is named “scientific knowledge map”. The scientific knowledge map can clearly reflect the development process and current situation of the research field. Compared with the traditional literature review, it is more systematic and intuitive, and has been applied in management, economics, psychology and other fields.

The literature samples in this paper come from Web of Science. In order to maintain the quality and value of the samples, the core journal database of Web of Science in the journal database is selected. The search keywords are “big data” and “organizational innovation”, and the time period is 2000–2021. In 2018, 726 papers were initially screened, and 579 samples were finally obtained after removing book chapters and conference proceedings papers.

3 Literature Quantity Trend Analysis

The overall number of publications reflects the attention, development level and overall evolution of the research topic. Through the statistics of the number of papers within a certain period, the overall trend of related research can be depicted. As shown in Fig. 1, the research on organizational innovation in the era of big data from 2000 to 2021 can be divided into three stages. The period from 2000 to 2011 was the initial stage. During this period, the research topic was not a hot spot in academia, and the number of papers published each year was less than 10. Among them, there were 0 papers in 2002 and 2003, and 6 papers in 2010 and 2011. The period from 2012 to 2015 was a

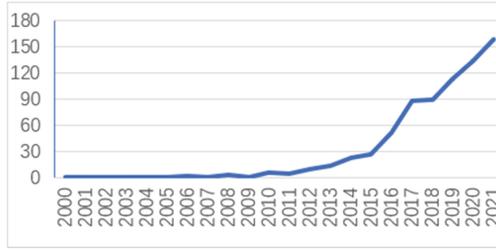


Fig. 1. The trend of the number of documents

period of slow development, during which the research topics gradually attracted the attention of researchers, and the number of published papers increased steadily year by year. The number of posts has nearly tripled over the years. From 2016 to 2021, the combination of big data and organizational innovation has attracted widespread attention from researchers. The number of published papers has increased rapidly year by year. In 2019, the number of published papers exceeded 100, and the number of published papers reached its peak in 2021, at 159.

4 Core Author Analysis

The research on the core authors in this paper is carried out in two parts. The first part analyzes the number of papers published by authors and the centrality of betweenness. Table 1 selects the top ten authors with the number of papers published. SABJAY KUMAR SIGH published 3 papers in 2019, with the largest number of papers, CASEY G Eight authors including CEGIELSKI, WAN KHAIRUZZAMAN WAN ISMAIL published 2 papers from 2017 to 2021, ranking 2nd to 10th. In the selected 579 samples, the number of papers published by each author is less than 3 papers. It can be seen that the distribution of organizational innovation research under the background of big data is relatively scattered. Table 1 shows the author's betweenness centrality. The top ten are JEFFREY MORGAN, MSTTHEW WILLIAMS, etc. The betweenness centrality is 9, and the distribution year is 2014. It can be seen that 2014 is an important year for this research topic. The papers published in this year enjoy a high status in academia. The second part is the analysis of the author's cooperation network, as shown in Fig. 2. The author's cooperation network has formed a main network and many sub-networks. The main network involves ten authors such as ROB PROCTER and ANITA GREENHLL. The network formed by these ten authors is complex. The number of cooperation, close contact, extensive exchanges, full communication, is conducive to resource sharing and optimization. In addition, ALVARO HERMAN GALVIS, ANGELICA AVALO, ALEXANDRA RAMIREZ team, JIABAO LIN, XIN(ROBERT) LUO, LEI LI team, YOGESH DWIVEDI, SHESHADRI CHATTERJEE, ALKIS THRASSOU team are also important research teams in the cooperative network.



Fig. 2. Author cooperation network diagram

Table 1. Author citation frequency and betweenness centrality

Counts	Authors	Centrality	Authors
3	SANJAY KUMAR SINGH	9	JEFFREY MORGAN
2	CASEYG CEGIELSK	9	MATTHEW WILLIAMS
2	WAN KHAIRUZZAMAN WAN ISMAIL	9	LUKE SLOAN
2	LEI LI	9	PETER BURNAP
2	YOGESH K DWIVEDI	9	ROB PROCTER
2	JIABAO LIN	9	WILLIAM HOUSLEY
2	SHIVAM GUPTA	9	OMER RANA
2	MANLIO DEL GIUDICE	9	ADAM EDWARDS
2	ABDULNASSER ELKASSAR	9	ALEX VOSS
2	SHESHADRI CHATTERJEE	9	ANITA GREENHILL

5 Keyword Frequency and Co-occurrence Analysis

Keyword is an accurate summary of the content and theme of the article. Keyword frequency reflects the attention of researchers in the field. Analyzing keywords can more accurately observe the hotspots in the research field. Using the CiteSpace analysis tool, this paper extracts a total of 345 keywords from 579 samples, and merges similar keywords, for example, “industry 4” is merged into “industry 4.0”, “organizational performance”, “firm performance”, and “performance” are merged into “performance” and so on, and finally extracted the top 20 keywords in frequency. As shown in Table 2, “big data” and “innovation” have the highest word frequency, 127 times and 120 times respectively. The search rules of the topic, followed by the word frequency of “performance” is 71 times, and the word frequency of “impact” and “big data analytics” has reached 50 times.

On the basis of word frequency analysis, in order to better understand the research hotspots of organizational innovation in the context of big data [2], it is also necessary to conduct keyword co-occurrence analysis. Keyword co-occurrence analysis mainly

Table 2. Top ten keywords with word frequency

Frequency	Keywords
127	big data
120	innovation
71	organizational performance
50	impact
50	big data analytics
47	management
45	information technology
35	capability
30	model
28	knowledge management

uses the co-occurrence of keywords in the same document to determine the relationship between the topics represented by the document set. It is generally believed that the more times the keywords appear in the same article, the more the closer the contact is. As shown in Fig. 3, the analysis unit is set as the keyword, the time span is from 2001 to 2020, the time slice is set to 1, and the network pruning method is minimum spanning tree, pruning slice network, and pruning merged network. The five largest nodes in Fig. 3 are “big data”, “innovation”, “impact”, “information technology”, and “knowledge management”. In the co-occurrence analysis, the size of the network nodes reflects the frequency of word occurrence, and the connection line indicates the co-occurrence relationship between the two. Based on the analysis of word frequency and co-occurrence, it can be seen that the most important keywords for organizational innovation research under the background of big data in the past two decades are information technology, knowledge management, business management, etc. The main research direction is the impact of organizational innovation on organizational performance under the background of big data. Impact, including innovation performance, management performance, financial performance, etc. [5].

6 Keyword Clustering Analysis

In order to more accurately understand the research themes of organizational innovation under the background of big data, and to mine the co-occurrence relationship between the themes of articles, this paper conducts a clustering analysis on the basis of the co-occurrence of keywords, and selects the LLR algorithm to obtain the clustering of academic research topics, as shown in Table 3. The modular Q value is 0.7512, and it is generally considered that a Q value greater than 0.3 is significant. The average contour value S of the contour is 0.8836. It is generally considered that the result is reasonable when the S value is greater than 0.5, and the result is convincing when the S value is greater than 0.7. It can be seen that the keyword clustering results in this



Fig. 3. Keyword co-occurrence diagram

paper are significant and have a certain degree of confidence. As shown in Table 3, the keyword clustering in this paper finally obtained 7 clustering labels, namely “organizational performance”, “supply chain management”, “knowledge management”, “knowledge-sharing behavior”, etc.

The #1 cluster is “organizational performance”, and the size is the largest, which is 69. Keywords included in this cluster include digital transformation, public sector innovation, scm performance, sustainable performance, etc. It can be seen that the research on organizational innovation in the context of big data mainly takes organizational performance as the starting point, focusing on the digital transformation of organizations, organizational innovation performance, supply Chain management efficiency and the sustainability of efficiency.

Cluster #2 is supply chain management. The researchers focus on the application of big data technology in the process of supply chain management, and study the green innovation of organizations in the field of supply chain. The keywords included in this cluster are big data analytics, green innovation, competitive advantage, manufacturing industry. With the development of big data technology, big data service providers gradually appear in the supply chain. The addition of big data technology reshapes the original supply chain decision-making method, forms new competitive advantages, and promotes green innovation and the development of manufacturing. At present, we mainly focus on two issues. The first is the cooperation between big data service enterprises as a separate subject in the supply chain, such as the coordination of energy conservation and emission reduction in the supply chain of big data service enterprises. The second is the application of big data technology in supply chain decision-making and coordination, such as the impact of information technology innovation on supply chain decision-making [1].

Cluster #3 and #4 are “knowledge management” and “knowledge-sharing behavior” respectively, and the keywords included are “knowledge sharing”, “innovative performance”, “intelligence” and so on. The accumulation of knowledge is a necessary condition for innovation, and knowledge management is an organization’s management of knowledge assets. With its powerful collection, storage and analysis capabilities, big data is a powerful tool for knowledge management. At present, there are two research

Table 3. Keyword Clustering

number	Size	Cluster labels
1	69	organizational performance
2	37	supply chain management
3	28	knowledge management
4	27	knowledge-sharing behavior
5	25	digital transformation

hotspots in academia. The first is the impact of big data on the transformation of knowledge sharing models, and then discusses the impact of big data on organizational innovation models. The second is the impact of big data on innovation performance, including financial performance, social value, etc.

Cluster #5 is “digital transformation”, which contains keywords such as “organizational learning”, “non-financial performance”, “knowledge management capability”, etc. It mainly studies how big data can promote the digital transformation of organizations. The main direction To leverage big data to facilitate organizational learning, break free from financial performance constraints, and improve knowledge management capabilities.

7 Keyword Mutation Analysis

Keyword emergent research on keywords with a sudden increase in the number of occurrences in a certain period of time indicates the degree of attention that the keyword received in the academic community at that time, and to explore potential research problems in the research field. Figure 4 reveals three aspects of enterprise innovation research under the background of big data. a stage. The first stage is from 2000 to 2011, during which big data technology is in its infancy, the number of keywords is small, and related research is in the initial stage, mainly to discuss the impact of big data technology on a certain industry and a certain type of managers, the related research is large and large, and has not been discussed in depth. The second stage was from 2012 to 2014. The related research was in a slow development stage, and the research hotspots were gradually refined, and began to involve keywords such as knowledge management, productivity, knowledge, etc. The mutation intensity of keyword productivity was 3.05, which was significantly higher than that in this stage. Other keywords, indicating that the biggest research focus at this stage is the impact of organizational innovation on organizational productivity in the context of big data. The third stage is from 2015 to 2021. This stage is a stage of rapid development of research topics, and the number of mutation keywords has increased sharply, involving “science”, “management”, “transformation”, “predictive analytics”, “cloud computing”, “supply chain management”, it can be found that there are various types of keywords involved, and research hotspots have begun to diverge. Among them, the keyword “science” has the highest mutation intensity at 3.39, indicating that researchers use data as a kind of science in organizational innovation.

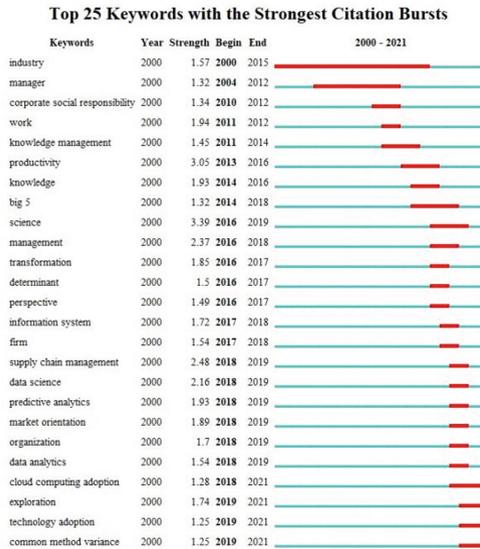


Fig. 4. Keyword mutation

The emergence of the theory indicates that the research on organizational innovation in the context of big data is maturing.

8 Conclusion

This paper takes the core journals of web of science as the data set, selects 579 samples, uses the citespace 5.8. R3 software to visually analyze the organizational innovation literature under the background of big data, and generates the overall trend map of the number of literature, the relationship map of research institutions, and the distribution map of core authors, keyword co-occurrence, clustering and mutation graphs. The findings are as follows. First, the research on organizational innovation under the background of big data can be divided into three stages. From 2000 to 2011 as the initial stage, the number of published papers was relatively small, mainly studying the impact of big data background on industry and the influence of managers; 2012–2014 is a slow development stage, and the number of documents has gradually increased, mainly to study the impact of organizational innovation on organizational productivity under the background of big data; 2014–2021 is a rapid development stage, in which the number of published papers has surged, the research topics have begun to diverge, and the field of research is maturing. Second, from the perspective of authors, the number of papers published by a single author is relatively low, generally less than 3 papers, and authors with a high number of papers have low intermediary centrality, indicating that related research is relatively scattered [3]. And the quality of the literature needs to be improved. Third, the research hotspot in the field of organizational innovation under the background of big data from 2000 to 2021 is the organic combination of information technology, knowledge management, enterprise management, etc. The main focus of the

research field is organizational performance, including innovation performance, management performance, financial performance, etc. Future research can be carried out in two directions. First, the influencing factors of organizational innovation in the context of big data, especially the influence of policy factors in the context of China, should be further explored. Secondly, it focuses on the dynamics of organizational innovation in the context of big data, and discusses what strategies organizations adopt in different stages.

References

1. CHENG Yanping, HUANG Qianzhi, DONG Ciwei. The application of big data in supply chain management: knowledge mapping analysis based on CiteSpace [J]. Journal of Central China Normal University(Natural Sciences), 2021, 55(03):453–461.
2. HE Qin. Research on Innovation Management of Flexible Employment People in Sharing Platform Organization under Big Data Driven [J]. Journal of Beijing Union University (Humanities and Social Sciences), 2019, 17(01):78–84.
3. QIU Xiping. Research on Organizational Structure Innovation of Data-empowering Manufacturing Enterprises [J]. Enterprise Economy, 2022, 41(01):84–93.
4. WANG Wan, CAI Wenwen, LI Runzhi. Research on the Influencing Factors of Enterprise Organizational Innovation in the Era of Big Data:Based on Combination Weighting Method and ISM Analysis [J]. Modernization of Management, 2020, 40(06):32–35.
5. YU Donghua, LI Yunhan. The Innovation of Industrial Organization in the Era of Digital Economy: Research on Industrial Chain-Cluster Ecosystem Driven by Digital Technology [J]. Reform, 2021(07):24–43.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

