

Development Trend of Data Protection: a Bibliometric Analysis and Review

Jinqun Luo, Yuefeng Yan

School of Cultural Communication and Design, Guangdong Open University, China

Abstract: Through the research on CNKI database, this paper uses bibliometrics and visualization tools to mine and statistically analyze the literature contents related to data protection and privacy security, summarizes the research status of data protection and privacy security, provides reference for subsequent research, and prospects the future research direction.

Keywords: Data protection; Privacy security; Visualization tools; Quantitative analysis

1. Introduction

In recent years, “data protection” has been mentioned more and more. The main reason is that there are news events about “data protection” both at home and abroad. In 2014, the iCloud server of Apple was exploited by hackers to steal the privacy data of Apple users. In 2021, China Didi travel company was accused of illegally collecting customers’ personal information. Data protection and personal privacy security have become issues of widespread concern to the public.

This paper analyzes the research results related to “data protection” and other similar topics in Chinese journal papers, summarizes the current situation, finds problems, and provides references for subsequent research directions.

2. Research Methodology

2.1 Data collection procedure

The research data were retrieved from the database of CNKI. The subject search terms are “data protection” and “privacy”, the time span is 2012-2021, and the source journal is set as “core journal”. Through database retrieval, 768 articles were retrieved. After analysis, excluding the documents about conferences, newspapers and periodicals that are less relevant to the theme, 740 documents were obtained, and the research data set was established.

2.2 Data analysis

This paper uses Bibliometrics method and visual analysis software VOSviewer to carry out graphic visual analysis. Bibliometric analysis refers to the cross science of analyzing all knowledge carriers by using mathematical and statistical methods and combining quantitative and qualitative analysis methods. VOSviewer is a tool for measuring and visualizing the literature metrology network. It can show the structure, evolution, cooperation, and other relationships in the field of knowledge, and help us quickly understand the development context and cutting-edge hot spots of relevant academic fields. This paper mainly studies the following issues:

(1) The increasing trend of literature publication.(2) Journal distribution.(3) Analyze the co-occurrence of keyword data to study the hot spots in this field.(4) Look for core authors and relevant core teams in this field.

3. Results analysis

3.1 Results about Literature output and growth trends

Research data shows that the number of studies on privacy data protection was only 12 in 2012, and has been growing slowly since then. The main reason is that scholars did not pay enough attention to the problem of privacy data protection at that time. Since 2016, the number of publications has increased year by year until it reaches 164 in 2021. This shows that the issue of privacy data protection has gradually attracted the attention of scholars, and the number of documents is increasing.

3.2 Results about distribution of Journals

From 2012 to 2021, these 740 articles were mainly published in two types of journals. The first category is legal journals, accounting for 44.86%, a total of 332, which shows that scholars have conducted a lot of research on the legislation of privacy data protection. The second category is computer journals, accounting for 40%, a total of 296, indicating that researchers have also conducted a lot of research on technical issues of data protection. The number of articles published in other types of journals is relatively small, which indicates that there is a trend of interdisciplinary research in privacy data protection, but the depth of interdisciplinary research is not high, and there is a lack of research from different perspectives. Future research will pay more attention to interdisciplinary and diversified in-depth research.

3.3 Results about most-used keywords

Among the 1720 keywords used, 43 have reached the threshold of using at least “9 times”, becoming the most commonly used keywords.

The hot content is mainly divided into three categories, namely, taking “privacy protection”, “big data” and “personal information” as the research directions, supplemented by intermediate frequency keywords such as “privacy, data security and cloud computing”. This shows that the research related to privacy data protection includes both the connotation of data protection itself and the technologies involved. At the same time, the research scope also includes legislation, data governance and cross-border data flow.

Among the 43 keywords, “privacy protection” is the word with the highest frequency, which is 246, indicating that it is one of the hot spots; The frequency of “big data” is 109, far ahead of other keywords, which shows that with the advent of the era of big data, privacy protection has attracted more public attention. Everyone is very concerned about the protection of personal privacy data, which shows that scholars do not pay attention to the protection of public data.

In terms of time weight, cross-border data flow, artificial intelligence, data governance and blockchain are becoming new hotspots.

3.4 Results about core authors and their relationships

The research results cannot be separated from the support of core authors. Through the statistical analysis of the comprehensive publication volume and citation frequency indicators of literature authors, core authors can be obtained. According to the two criteria of being published more than 3 times as the first author and being cited more than 25 times, 17 core authors were obtained.

The results show that 15 of the core authors are from the legal category, including Gao Fuping^[1], Ding Xiaodong^[2], Huang Ruhua^[3] and Gong Yongqin^[4], which shows that Chinese scholars attach great importance to the legislation of privacy data and have formed representatives.

Another major source of periodicals is computer. But there are only two core authors, Ma Xiaoting^[5] and Wang Taochun^[6]. The data shows that although computer researchers have done a lot of research, they have not produced many representatives.

In order to further study the cooperative relationship between the core authors, this paper uses VOSviewer software to set the minimum number of posts to 2, and 177 qualified authors can be obtained. Research data shows that the research author group is scattered, but also formed a certain core team.

At present, the largest team is Wang Taochun’s 18-member team, mainly from Anhui Normal University. The second is Huang Liuheng’s 8-member team, almost all from the University of science and technology of China. They are mainly engaged in computer science research.

It can be seen that legal scholars are used to individual research, while computer scholars are mainly team research.

Moreover, these teams mainly come from the same unit or discipline, and there is a lack of interdisciplinary or cross unit team cooperation research. Future research should pay more attention to interdisciplinary and cross unit teamwork research.

4. Conclusion

This paper makes a quantitative analysis of data protection and privacy information security research in China. Research findings:

4.1 Lack of interdisciplinary and interdisciplinary research teams.

The research results show that the current domestic research on “data protection” and “privacy security” lacks interdisciplinary and cross unit team cooperation research. Law and computer scholars have done a lot of research on “data protection” and “privacy security”, but law scholars are used to personal research, while computer scholars are mainly team research, even if they do team research, they almost come from the same unit. There is a lack of interdisciplinary research teams of law and computer scholars, as well as interdisciplinary team cooperation in other disciplines.

4.2 The protection of personal privacy information and public information needs more attention.

The research results show that there is not much domestic literature on privacy protection legislation. Scholars mainly learn from foreign general principles of data privacy protection, and have not formed a high-value and mature research system.

The disclosure of private data of individual users and public information of customers mentioned above have proved the importance of protecting personal information and public data security on Internet applications.

5. Suggestion

5.1 The emergence and development of new information technology also need to be included in the scope of research.

According to the research results, cross-border data flow, artificial intelligence, data governance and blockchain are becoming new hotspots. This reflects that in the era of new information technology change, the development of new technology needs continuous improvement of laws and regulations to govern and respond.

The importance of data protection and personal information security has become increasingly prominent, and the relevant supervision of national laws and regulations also needs to be continuously improved.

5.2 For data governance, we need to strengthen the cooperation between computer teams and law teams.

It is particularly necessary to carry out interdisciplinary research team cooperation, especially in computer and law teams. Computer teams can evaluate and summarize the risks of data protection and information security. Law teams can standardize and manage the risks of new technologies based on the revision of existing laws and regulations, so as to achieve a relatively stable and safe dynamic balance between technological development and legal perfection.

References:

-
- [1] Gao Fuping Personal information protection: from personal control to social control [J] Legal research, 2018, 40 (3): 18.
 - [2] Ding Xiaodong The dilemma and outlet of private law protection of personal information [J] Legal research, 2018, 40 (6): 13.
 - [3] Huang Ruhua, Liu long Problems and Countermeasures of personal privacy protection in the opening of government data in China [J] Library, 2017 (10): 5.
 - [4] Gong Yongqin, Wang Jian Research on the rules of cross-border flow of personal data between APEC and EU [J] Asia Pacific economy, 2015 (5): 5.
 - [5] Maxiaoting Research on library personalized service and reader privacy protection in the era of big data [J] Library forum, 2014, 34 (2): 6.
 - [6] Wang taochun, Liu Ying, Jin Xin, et al Research on location and data privacy protection methods based on k-anonymity in group intelligence perception [J] Journal of communications, 2018, 39 (A01): 9.

* **Corresponding author:** Jinqun Luo