Research on Frontier Issues of Privacy from the Perspective of Knowledge Map

Xiaoxiao Zhan¹,a

¹Anhui University of Finance and Economic, Bengbu, China

a1161679582@qq.com

Abstract: Taking 1591 core journal documents included in CNKI and wos databases as the data sample source, using research methods such as co-word analysis, keyword analysis and time zone analysis, and using visual analysis tools such as city space, this paper makes an in-depth analysis of the research on privacy protection at home and abroad in recent 10 years. This paper summarizes and analyzes the structural characteristics, development context and hot frontier of privacy research at home and abroad from the macro and micro levels. The research shows that the right to privacy and privacy protection are interdisciplinary and complex research, and their research fields focus on computer software and application, civil and commercial law and administrative law, news and media, etc; Research hotspots mainly focus on technology, law and application market. Frontier topics include blockchain, location privacy, the right to be forgotten, privacy concerns, information security, etc; The research in the field of privacy has roughly experienced the main development path of "privacy construction and social network → network topics include blockchain, location privacy, the right to be forgotten, privacy concerns, information security, etc; The research in the field of privacy has gradually occupied an important place in the legal field. At the same time, the development of information technology makes privacy inseparable from personal information, Internet, e-commerce, big data, cloud computing and artificial intelligence. At present, the research on the right to privacy at home and abroad is mainly carried out from the perspective of qualitative analysis and logical demonstration. For example, Lu Bingwan qualitatively demonstrates the dilemma and Countermeasures of the right to privacy of information outside the territory from the perspective of personal information control, and comes to the conclusion that the tendency of privacy protection in China under the epidemic is the civil rights that can be transferred to the national and collective interests [1]. Ren Ying proposed to reconstruct the legal structure and rules of privacy protection in the data age under the requirements of the principle of balance [2]. Through comparative analysis, Ni Fei demonstrated the advantages of competition law in the field of consumer privacy protection, and proposed that it is necessary to improve legislation and make up for the institutional shortcomings involved in privacy protection [3].

The analysis of privacy research literature is mainly carried out by qualitative comparison method, and the research of quantitative analysis literature is less. Ouyang Yang et al. Analyzed the research progress of privacy concern in e-commerce at home and abroad by taking the research literature in the field of e-commerce privacy concern at home and abroad as the research object[4]; Based on the perspective of mode transformation, Zhu Yue summarized the characteristics of existing legal protection and summarized the research on legal protection of personal information under the background of big data from the perspective of defining the concept of personal information [5]; Li Zhu compared and analyzed 30 literatures from 12 to 19 years [6], and discussed the privacy protection literature in digital life from the perspective of research comparison and prospect at home and abroad; Xu Hanwen and others reviewed the research on data rights in the United States by means of bibliometric analysis [7].

1. Introduction

The development of the Internet has brought qualitative changes to people's life. The network has gradually become an essential component of people's life, so that people can enjoy a variety of convenient and fast services without leaving home. At the same time, the advent of artificial intelligence and 5g commercial era also brings the infringement of the rights and interests of traditional physical space to cyberspace. Massive data collection and analysis has challenged the protection of citizens' personal privacy. Whether it is web browsing traces, shopping consumption records, account privacy such as mobile phone number and credit card number, location privacy such as address and workplace, and communication privacy such as video records and voice records, they are monitored by big data. Privacy issues have aroused widespread concern.

According to the statistics of the United Nations Trade Development Organization, as of February 21, 2022, about 80% of the world's countries (194 countries in total) have completed data security and privacy legislation or proposed draft laws, including EU Member States, the United States, China, Russia, India, Australia, Canada and Japan. With the deepening of digital transformation, the situation of data security and privacy is becoming more and more serious. Modern data security and privacy protection legislation has become a global trend. Today, with the frequent occurrence of network privacy events, privacy has become the focus of both industry and academia.

2. Research Status and Literature Review

Since American lawyers Warren and Brandis jointly published the article "the right to privacy" in 1890, the right to privacy has gradually occupied an important place in the legal field. At the same time, the development of information technology makes privacy inseparable from personal information, Internet, e-commerce, big data, cloud computing and artificial intelligence. At present, the research on the right to privacy at home and abroad is mainly carried out from the perspective of qualitative analysis and logical demonstration. For example, Lu Bingwan qualitatively demonstrates the dilemma and Countermeasures of the right to privacy of information outside the territory from the perspective of personal information control, and comes to the conclusion that the tendency of privacy protection in China under the epidemic is the civil rights that can be transferred to the national and collective interests [1]. Ren Ying proposed to reconstruct the legal structure and rules of privacy protection in the data age under the requirements of the principle of balance [2]. Through comparative analysis, Ni Fei demonstrated the advantages of competition law in the field of consumer privacy protection, and proposed that it is necessary to improve legislation and make up for the institutional shortcomings involved in privacy protection [3].

The analysis of privacy research literature is mainly carried out by qualitative comparison method, and the research of quantitative analysis literature is less. Ouyang Yang et al. Analyzed the research progress of privacy concern in e-commerce at home and abroad by taking the research literature in the field of e-commerce privacy concern at home and abroad as the research object[4]; Based on the perspective of mode transformation, Zhu Yue summarized the characteristics of existing legal protection and summarized the research on legal protection of personal information under the background of big data from the perspective of defining the concept of personal information [5]; Li Zhu compared and analyzed 30 literatures from 12 to 19 years [6], and discussed the privacy protection literature in digital life from the perspective of research comparison and prospect at home and abroad; Xu Hanwen and others reviewed the research on data rights in the United States by means of bibliometric analysis [7].
To sum up, in the context of the deepening application of big data, looking at the domestic and international academic circles, although the literature achievements related to the right to privacy are quite fruitful, there are relatively few studies on literature combing and analysis from the perspective of quantitative analysis. Visual analysis with the help of analysis methods such as knowledge map can make up for the lack of this context research, which is of great significance.

3. Data Sources and Analysis Tools

3.1. Data Sources

The research direction of this study is to explore the frontier issues of the right to privacy, and the research object is limited to the journal literature of the right to privacy and related concepts. Considering that privacy is a concept involving many disciplines, including information science, management, social psychology, law, economics, computer science and other disciplines have explained privacy from the dimension of their own disciplines, therefore, in the selection of journal papers, the whole science and technology contribution is taken as the research object to ensure the accuracy and comprehensiveness of the research.

Through the retrieval of CNKI and WOS databases, it is found that from 2012 to 2022, the number of research literature related to privacy reached 11843, and only 3173 core journals and CSSCI journals were searched. If the traditional literature reading method is adopted, it is difficult to comprehensively summarize the key literature and research hotspots in a certain field, and researchers are easy to fall into the trap of single subjective judgment. The appearance of visual econometric analysis tools such as CiteSpace and VOSviewer simplifies the cumbersome traditional literature browsing form, and combs the literature from the perspective of the combination of subjective judgment and objective measurement.

Therefore, this study specifically takes the CNKI database literature from 2012 to 2022 as the data source, collects the relevant journals of domestic and foreign scholars on the research of privacy, the source categories are SCI source journals, core journals and CSSCI, and carries out advanced retrieval with "privacy", "privacy protection", "network privacy", "personal information protection" and "privacy paradox", the matching method is "exact matching", the journal type is limited to "core journal", "CSSCI" and "SCI", and the journal year is set from 2012 to now (the search time is March 23, 2022). Other search conditions are set as the system default. A total of 1591 journal documents were obtained, and then they were checked manually one by one, excluding obviously irrelevant documents such as announcements, statements and draft notices. The final valid titles were saved in Refworks data format and downloaded as the actual data source of this study.

The CiteSpace tool contains two parts of parameter settings: initial interface parameter settings and network visualization and editing interface parameter settings after map generation. The initial interface parameter settings include time zone segmentation, network setting, threshold setting, network simplification and other setting buttons. The actual data source is imported into CiteSpace analysis tool for data statistics and generation of scientific knowledge map, and the relevant parameters are set. Take typical parameters as an example:

"Term source": used to select the extraction location of subject words, including title, abstract, author keywords and keywords plus.

"Node types": author, term, keyword.

"Topn" (frequency selection): n = 30 (select the keywords and subject words that appear in the top 30 every year).

"Pruning": Pathfinder.

According to the different analysis contents, the research selects specific keywords, draws the relevant knowledge map, and makes a systematic analysis. At the same time, in the process of drawing, in order to make the map scientific and readable, some parameters are adjusted appropriately, so as to summarize and sort out the knowledge structure, development context and cutting-edge trend of privacy research at home and abroad from the macro and micro level.

3.3. Parameter Adjustment

By searching CNKI and WOS database documents from 2012 to 2022, the search conditions are: subject / Title = "privacy", "privacy", "privacy protection", "network privacy", "personal information protection" and "privacy paradox", the matching method is "exact matching", the journal type is limited to "core journal", "CSSCI" and "SCI", and the journal year is set from 2012 to now (the search time is March 23, 2022). Other search conditions are set as the system default. A total of 1591 journal documents were obtained, and they were checked manually one by one, excluding obviously irrelevant documents such as announcements, statements and draft notices. The final valid titles were saved in Refworks data format and downloaded as the actual data source of this study.

3.2. Analysis Tools

Combined with the data involved in this study and the research purpose, this paper selects the CiteSpace visual analysis software developed by Professor Chen Chaomei as the analysis tool to analyze the research hotspots, research frontiers and research trends of privacy from 2012 to 2022 according to the visual map [8] obtained after data processing.

4. Visual Analysis of Privacy Research

4.1. Annual Document Information Statistics

Using CNKI visualization tool, search the "privacy" literature in recent ten years, classify and summarize the publication year of 1591 documents obtained in CNKI database, and obtain the statistical chart of annual document information shown in Figure 1.
Figure 1. Annual Distribution of Research Literature on the Theme of "Privacy"

As can be seen from Figure 1, the trend of papers on privacy protection research in recent ten years is generally flat and upward, which shows that after the accumulation of research results and research methods since 2012, researchers in this field can quickly and sensitively find the problems to be studied and find the methods to deal with them properly, so as to publish research results faster and promote the transformation and landing of results. This promotes the growth of the number of articles. At the same time, the research focus on the right to privacy continues, the occurrence of infringement events such as Facebook privacy disclosure in 18 years [9], and major events such as the first determination of the right to privacy in the civil code in 2021 promote the improvement of the protection of the right to privacy and the development of relevant research.

4.2. Subject Distribution

Using CNKI visualization tool, we can get the trend of analysis and subject analysis, as shown in Figure 2.

Figure 2 Discipline Distribution of Documents on the Theme of "Privacy"

As can be seen from Figure 2, in recent ten years, the literature related to the theme of privacy has been mainly distributed in civil and commercial law, administrative law and local legal system, news and media, procedural law and judicial system, law and legal history, economic law, constitution, etc. At the same time, according to the descending order of the number of research results in different fields, it is found that the results produced by the top 10 fields account for more than 80% of the total, indicating that the top 10 fields are in the first echelon of privacy research. This also includes the fields of "news and media", "computer software and computer application" and "archives and museums", which reflects the obvious interdisciplinary attribute of privacy research.

News and media mainly carry out research from the balance between information disclosure and information security. The high-speed dissemination of information greatly promotes the work of information disclosure and enables the public to obtain information quickly. However, due to the characteristics of information dissemination, the control of information is weak and vulnerable to infringement. Therefore, the research mainly focuses on the balance between information disclosure and information security. From the perspective of computer software and computer application, it mainly studies the data mining, acquisition and protection technology of personal information. In view of the large output and capacity of information in the era of big data, all kinds of personal information are exposed on the Internet. Therefore, in the computer science, it mainly studies how to reduce users' disclosure of personal sensitive information, so that personal information can be legally obtained and used by the society, so as to improve the quality of life, and how to protect the information after users inadvertently disclose it, such as using de identification, anonymization and other technologies to process personal information. The discipline of archives and museums mainly studies the types of personal information and the ways of disclosure. Users' browsing traces such as account information, behavior information and basic information involved in digital library can reflect users' interests and reading habits, which is convenient for the system to make relevant recommendations for users. However, if there is leakage or malicious attack in data
management, the rights and interests of readers will be infringed. Reasonable use of readers' personal information while maintaining the security of their information has become a research hotspot, which needs the joint maintenance of legal construction, self-awareness, network technology, industry self-discipline and so on.

4.3. Periodical Distribution

It can be found from table 1 that the journals with high volume of papers in the field of personal information protection in China are mainly distributed in the discipline fields relatively close to information protection research, such as information and information, politics and law, computer technology, information and knowledge dissemination, scientific research and so on. High frequency journals such as information journal, information theory and events pay attention to the technical research of privacy protection, so researchers can pay attention to these journals when conducting relevant research. The research journals in the field of politics and law mainly include law and business research, Chinese law and emotional law research, which mainly focus on the possible legal regulation problems in the process of dissemination and preservation of personal information. At the same time, there is a high frequency of documents in the field of computer technology and the field of information and knowledge dissemination, which is enough to see that privacy protection is a multidisciplinary research hotspot that has attracted extensive attention. The above journals have high comprehensive impact factors and composite impact factors, which have a great impact in China. They have conducted multangle research on privacy protection, which provides a basis and guiding direction for researchers to continue to study related issues.

<table>
<thead>
<tr>
<th>Title</th>
<th>Number of articles published (articles)</th>
<th>Proportion of each journal (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intelligence magazine</td>
<td>16</td>
<td>13.11%</td>
</tr>
<tr>
<td>Information theory and events</td>
<td>16</td>
<td>13.11%</td>
</tr>
<tr>
<td>Legal and commercial research</td>
<td>11</td>
<td>9.02%</td>
</tr>
<tr>
<td>Chinese law</td>
<td>8</td>
<td>6.56%</td>
</tr>
<tr>
<td>Legal research</td>
<td>8</td>
<td>6.56%</td>
</tr>
<tr>
<td>Zhejiang Social Sciences</td>
<td>7</td>
<td>5.74%</td>
</tr>
<tr>
<td>Library and information work</td>
<td>6</td>
<td>4.92%</td>
</tr>
<tr>
<td>Law</td>
<td>6</td>
<td>4.92%</td>
</tr>
<tr>
<td>Exploration and contention</td>
<td>5</td>
<td>4.10%</td>
</tr>
<tr>
<td>Chinese and foreign law</td>
<td>5</td>
<td>4.10%</td>
</tr>
<tr>
<td>Journal of Wuhan University</td>
<td>4</td>
<td>3.28%</td>
</tr>
<tr>
<td>Nanjing Social Sciences</td>
<td>4</td>
<td>3.28%</td>
</tr>
<tr>
<td>Academic Monthly</td>
<td>4</td>
<td>3.28%</td>
</tr>
<tr>
<td>Henan Normal University</td>
<td>4</td>
<td>3.28%</td>
</tr>
<tr>
<td>Library and information</td>
<td>4</td>
<td>3.28%</td>
</tr>
<tr>
<td>Chinese Social Sciences</td>
<td>3</td>
<td>2.46%</td>
</tr>
<tr>
<td>Information Science</td>
<td>3</td>
<td>2.46%</td>
</tr>
<tr>
<td>social sciences</td>
<td>3</td>
<td>2.46%</td>
</tr>
<tr>
<td>Library Forum</td>
<td>3</td>
<td>2.46%</td>
</tr>
<tr>
<td>Chinese Translators Journal</td>
<td>2</td>
<td>1.64%</td>
</tr>
</tbody>
</table>

5. Research Hotspot Analysis

5.1. Analysis of Research Hotspots on the Right to Privacy

In order to analyze the research hotspots of privacy protection, this study uses the data samples determined above to draw the knowledge structure map of the research field of "privacy and privacy protection", and analyze the theoretical structure and research hotspots of this research field at home and abroad in recent 10 years. The specific operation methods are as follows: set time SLI cing in CiteSpace to 2012-2022, node types to key word and selection criteria to T50 to obtain the knowledge map of domestic research hotspots, as shown in Figure 3. The keyword is automatically generated according to the built-in database. The larger the circle diameter and font size of each node, the more research results of the keyword; The darker the color of the node, the earlier the research results of this topic are obtained within the research time range of this paper; The thicker the connection between fields, the closer the cooperation relationship between the two topics and the more results of cooperation.

As can be seen from Figure 3, "privacy protection", "right to privacy", "location privacy", "personal information", "social network" and other keywords have high frequency and centrality, indicating that the research always focuses on privacy protection and right to privacy, and makes peripheral expansion research in relevant legislation and technology.

It shows that the use, management and protection of personal information and privacy have gradually become a topic of common concern, and individuals, the state and society pay full attention to it from all aspects. With the advent of the era of big data in recent years, many personal information on the network has been widely excavated and exposed, and citizens' privacy has been threatened. From the literature analysis, scholars at home and abroad pay more attention to network supervision [10,11], electronic information protection [12 16], privacy concern [17,18], privacy paradox [19 22], privacy computing [23,24], etc., from data mining and preprocessing to the use of relevant technologies for information hiding, Reduce information...
leakage. From the research frequency and centrality of the "right to privacy", there are still deficiencies in the legal protection of personal information and privacy in China. Although the "right to privacy" system is clearly established in the civil code, the corresponding implementation measures are still missing. Relevant scholars are studying how to further formulate relevant norms to adapt to the characteristics of the times for our existing national conditions and existing relevant laws.

5.2. Co-citation Analysis

In CiteSpace, select the reference node and slice the time according to each year to obtain the co cited knowledge map of the privacy research conducted in this paper, as shown in Figure 2. The larger the diameter and font size of the node circle in the figure, the more times the document has been cited; The darker the color of the node, the earlier the publication time of the literature represented by the node; The thicker the connection between nodes, the higher the co citation frequency between the two literatures.

Figure 3 Map of Hot Knowledge in Privacy Research

Figure 4. Co-citation Map of Privacy
As can be seen from Figure 4, the color of most nodes is darker, that is, most of the cited documents are published relatively early, which shows that the international research on the privacy system has a good continuity and can play a good role in inheriting academic ideas [25 ~ 27]. Most of the documents that play a major role in the citation network are concentrated between 2015 and 2021. The co cited documents that play a major role are not only the basis for subsequent research in this field, but also can reveal the research trends and changes in related fields in a certain period of time. Through the analysis of the above documents in CNKI and WOS databases, it can be found that the above documents are related to privacy and other contents. This shows that the research on relevant systems runs through the international intellectual property research, and continues to play a corresponding role in this field. Among them, "data mining" and "privacy" are highly cited. They mainly focus on the fields of information and knowledge dissemination and political and legal research, and study the problems in information dissemination such as data mining and preprocessing. The research on the protection of the right to privacy, the right to know, the right to Internet privacy and other legal fields around the "right to privacy" not only shows a high degree of citation, but also has a long period of citation, indicating that the right to privacy has become the main research hotspot during this period [28 ~ 32]. Then the keyword "human flesh search" also attracted attention in the fields of politics, law and computer application and technology, and attracted scholars' legal research on the right of reputation and infringement of network privacy, as well as the technical protection measures of personal information [33,34].

5.3. Analysis on Research Hotspots of Privacy Protection

The keyword research hotspot time zone map can show the distribution and changes of keywords in a topic or field in different time periods, and help us grasp the hot topics and their evolution process in the field of privacy from the time dimension. Set the network node as the keyword, then run CiteSpace, select "time zone" in the layout panel, and get the visual map of the time zone of research hotspot changes from 2012 to 2022. Observing the time zone chart of key words in privacy research literature in recent 10 years (Figure 5), we can find that privacy is an eternal hot topic in privacy research, and its research evolution process can also be summarized as the change process of privacy protection perspective. It has roughly experienced the main development path of "privacy construction and social network → network privacy and big data → blockchain and track privacy". The research topics tend to be diverse and the research contents continue to deepen, Horizontal and vertical differentiation accelerated.

In the first stage (2012-2015), scholars at home and abroad paid attention to the discussion of the right to privacy. The key words in this period are mainly privacy protection, right to privacy, privacy policy, public figures, etc., mainly discussing the privacy rights that individuals should enjoy and how to protect privacy from the perspective of law or policy [35 ~ 40]. From 2012 to 2015, it mainly focused on the research on how to carry out data fusion, homomorphic encryption, cloud storage and other technical issues for mobile Internet and social networks, so as to reduce the possibility of information infringement while bringing information convenience to individuals and society.

In the second stage (2016-2018), the problem of network privacy further emerged. The keywords in this period are mainly K anonymity, secure multi-party computing, data mining, network privacy, homomorphic encryption, association rules, etc., focusing on how to use technical means to solve the problem of network privacy [41 ~ 44]. At the same time, the emergence of the word "big data" scholars focus on the threat of data characteristics in the era of big data to personal information [45 ~ 51], the regulation and exploration of personal information protection [52 ~ 60]. Affected by the emergence of big data, data sharing has increased, and Research on industry self-discipline and
privacy policies for personal information has increased. The third stage (2019-2022) mainly focuses on blockchain, track privacy and other aspects, which will continue. Through the application of blockchain technology, the data sovereignty of users can be returned to users, and the personal identity information of Internet users and online consumers can be protected. For the protection of track privacy, research is made to improve the protection effect, service quality and operation efficiency in terms of the number of tracks released and the security of LBS server [61 ~ 65]. Among them, the research focuses on social networks and location services. At the same time, with the development of wireless mobile devices, personal location is exposed at any time, and the problem of location privacy disclosure has attracted attention. The research analyzes its main security threats and solutions [66,67] and how to improve data availability while protecting data privacy [68 ~ 75].

6. Summary

Taking 1591 core journals in the field of privacy collected by CNKI in recent 10 years as the data source, and using CiteSpace and other visual analysis software, this study summarizes and analyzes the research in the field of privacy from three aspects: structural characteristics, development context and hot frontier, and draws the following conclusions:

Structural features: in recent 10 years, the research heat of privacy in China has been rising steadily. The research on the subject of privacy presents the characteristics of diversity and intersection, involving different fields and professional backgrounds. It has strong attribute characteristics of cross research. It is involved in the fields of law, politics, engineering, computing science, library and information. There are key documents in the research on the intellectual property system in all fields, which makes the research on the intellectual property system more comprehensive. The research results are more operable.

Hot Frontier: the hot and frontier topics of current privacy research include location privacy protection, the right to be forgotten, privacy concerns, differential privacy, Internet of things, cloud computing, social media and information security. The research frontiers and hotspots of personal information protection are reflected in technology, law and application market. Among them, "data mining" and "privacy" are highly cited. They mainly focus on the fields of information and knowledge dissemination and political and legal research, and study the problems in information dissemination such as data mining and preprocessing. Around the "right to privacy", the research in the legal fields of privacy protection, the right to know, Internet privacy and human flesh search has become the main research hotspot.

Development context: the research in the field of privacy has roughly experienced the main development path of "privacy construction and social network → network privacy and big data → blockchain and track privacy". The research topics tend to be diverse, the research content continues to deepen, and the horizontal and vertical differentiation accelerates. The future research trends in the field of privacy in China include the discussion of the right to privacy under the background of new technology, the privacy personalization paradox in the era of mobile Internet, and the privacy protection in the era of 5g network.

Acknowledgment

2022 Anhui University of Finance and Economics Undergraduate Research and Innovation Fund Project (Project No.: XSKY22047ZD).

References


[22] Cao Xiaojuan Redefinition of privacy under the "privacy paradox" of social media [J] Satellite TV and broadband multimedia, 2020 (05): 129-130.


[50] Chen Yuyong, Lu Zhipeng Network monopoly of Internet platform enterprises and the protection of citizens' privacy --


[58] Zhao Hong From information disclosure to information protection: the circulation and core issues of information right protection in public law [J] Comparative law research, 2017 (02): 31-46.


[60] Yin Le, Li Yi Privacy issues in Internet Governance: personal life sharing and privacy protection based on social media [J] News and communication research, 2016,23 (S1): 69-77 + 127.


[67] Liu Yanhong Legal logic and limits of large-scale monitoring in public space -- from the perspective of orderly sharing of personal information [J] Law forum, 2020,35 (02): 5-16.


