

Comparative Analysis on Intangible Cultural Heritage Digitization in Chinese and International Journals: A Bibliometric Perspective

Zhaozhao Hu, Lianfan Wu, Xiaowei Qian*

Wenzhou Polytechnic, Wenzhou 325035, P. R. China

Abstract

In recent years, with the rapid development of digital technology, digitization has become one of the main ways to protect and disseminate intangible cultural heritage (ICH). In this study, 298 papers in the Web of Science (WoS) database and 240 papers in the China National Knowledge Infrastructure (CNKI) database from January 2004 to December 2021 were used as samples to compare and analyze the research on digitization of ICH in China and abroad. By analyzing and visualizing the information of publication time, discipline distribution, country distribution, authors, institutions, and keywords, the results revealed that: (1) Most international research focuses on scientific and technical disciplines, and computer disciplines occupy an important position in the research of ICH digitization. Chinese research, for the most part, focuses on humanities disciplines, and the proportion of computer disciplines is lower than international levels. (2) In recent years, international research involves emerging disciplines such as green sustainable science technology, environmental sciences, environmental studies, remote sensing, and materials science multidisciplinary, but Chinese research has not yet covered the above-mentioned fields. China should strengthen the cross-fertilization between different disciplines, encourage more engineering disciplines to participate in research, and focus on the research and application of emerging technologies. (3) Whether in China or abroad, there is no stable core author group in the field of ICH digitization. There are few collaborations among authors. (4) In the future, China should stimulate the communication vitality of various research subjects, build an open and integrated innovation ecology, and jointly promote the construction of a collaborative research network for digitization of ICH.

Keywords

Intangible Cultural Heritage Digitization; Bibliometric Analysis; CiteSpace; Knowledge Map.

1. Introduction

In 2003, UNESCO adopted the "The Convention for the Safeguarding of the Intangible Cultural Heritage". Since then, intangible cultural heritage (ICH) has attracted worldwide attention, and various disciplines have gradually participated in the research on its protection, inheritance, dissemination and development [1]. Under the background of the establishment of the concept of ICH and the gradual maturity of digital technology, the digitization of ICH, as a specific form of ICH protection, has attracted the attention of scholars from all over the world. The research on digital protection of ICH mainly focuses on three aspects: digital collection, digital preservation and digital development [2]. Digital acquisition focuses on the compilation of digital inventory, digital production & reconstruction and digital recording & archiving; digital preservation focuses on the organization, retrieval and management of digital information; digital development focuses on digital display, database construction and digital information communication.

China joined the ICH Convention in 2004. As of April 2022, China has a total of 42 items included in the World Intangible Cultural Heritage List, ranking first in the world in total [3]. In recent years, China has paid more and more attention to the digitization of intangible cultural heritage. However, the current research on the digitization of ICH is mainly based on applied research, and there is a lack of systematic sorting out.

With the maturity of bibliometrics, bibliometric analysis methods and software have gradually expanded from research in the field of library and information to research in other disciplines [4], and have many applications in the field of intangible cultural heritage research. For example, Liu H, Liu Y [5], Chen C, Huang Y [6], and Duan X [7] conducted a knowledge map analysis of the research literature on Chinese ICH in the CNKI database in the past 20 years with CiteSpace, analyzing the distribution of publications, authors, institutionals, and keyword co-occurrence, and summarizing the current status of research on ICH archives in China in recent years. By comparing the differences in the participation of various disciplines and research hotspots in ICH research in China and globally, Zhang Q, and Wang D [1] found that the participation of computer science in ICH research is very high in the international context, and suggested that Chinese scholars should invest more in the digitization of ICH. In addition, some scholars have conducted bibliometric and visualization analyses of Chinese traditional sports, especially Chinese traditional martial arts, in recent years to summarize research hotspots and predict research trends [8-12]. Some scholars have also conducted bibliometric analysis on specific minority ICH studies [13-15].

However, judging from the existing research, there is a lack of review studies on the development of research on digitization of ICH in China and abroad. In this paper, relevant documents in the field of ICH digitization were retrieved from the Web of Science (WoS) database and the China National Knowledge Infrastructure (CNKI) database respectively, and the CiteSpace software and the principles and methods of knowledge map were used to present the research progress and evolution of digitization of ICH in China and foreign countries.

2. Research Methods and Study Data

2.1. Study Data

2.1.1. International Literature

Since WoS database provides more complete references, indexes, and researcher relations than other databases, it was used as the data source for this study. The advanced search function provided by the core collection of WoS database was used on April 4, 2022 to retrieve all documents with the topics of "Intangible Cultural Heritage" and "Digital" until December 2021. "The types of literature included articles, proceedings papers, early access and review articles. The search yielded 305 documents, which were manually screened for non-relevant documents, resulting in 298 relevant documents.

2.1.2. Chinese Literature

In this study, the Core Journals of Peking University and CSSCI databases in China National Knowledge Infrastructure (CNKI) were selected as the data sources of Chinese literature. Because CNKI is the most important academic database in China, and Peking University Core Journals and CSSCI database only list high-quality literature in CNKI. On April 4, 2022, through the advanced search function provided by CNKI database, we searched the topic "intangible cultural heritage" and "digitization", and manually screened non-relevant literature. Finally, we obtained 240 relevant Chinese literature.

2.2. Research Methods

CiteSpace is an important software in bibliometrics, and knowledge mapping is a tool used to identify and display research dynamics and predict research trends in literature research. This

study uses CiteSpace5.8.R3(64-bit) to organize and analyze the literature of ICH digitalization in China and abroad respectively, including its yearly distribution, discipline and country distribution, main research authors and institutions, keyword co-occurrence, keyword clustering, keyword highlighting, in order to sort out the differences between Chinese and foreign research in this area through the obtained graphs and data.

In this study Citespace are used to generate the co-occurrence network map with the following parameter settings: Due to the different publication dates of the first papers in the field of ICH digitization research in China and abroad, the time slice was set to "From2004 to 2021" when analyzing the WoS data and "From2006 to 2021" when analyzing the CNKI data, with the same other parameters. The Slice Length is set to 2, the Selection Criteria is set to Top N, and select top 50 levels of most cited or occurred items from each slice, and the links strength is calculated by cosine distance, i.e. [16]

$$\cos(\theta) = \frac{\sum_{i=1}^n (X_i \times Y_i)}{\sqrt{\sum_{i=1}^n (X_i)^2} \times \sqrt{\sum_{i=1}^n (Y_i)^2}}$$

This represents the data of the top 50 nodes with the highest frequency of citations or occurrences per 2 years in the data set in the range of time slicing years for the literature of ICH digitization research, and the results are obtained by log-likelihood ratio (LLR) cluster analysis. In addition, the knowledge graph pruning method used in this study is Pathfinder, which simplifies the relationships to measure data similarity in the complex network of the knowledge graph of "ICH digitization" by checking the relationships between all data and keeping only the strongest links among all possible two-point paths, so as to establish the the most efficient linking path between the hot data. Its calculation principle is based on the following triangle inequality,

$$w_{ij} \leq \left(\sum_k w^r n_k n_{k-1} \right)^{1/r}$$

where W_{ij} is the weight of a direct path between i and j , $wn_k n_{k+1}$ is the weight of a path between n_k and n_{k+1} , for $k = 1, 2, \dots, m$. In particular, $i = n_1$ and $j = n_k$. [17]

3. Study on the Distribution of ICH Digitization Fields

3.1. Yearly Distribution of Publications

Analysis of the distribution statistics of the literature publication time can help us understand the academic progress and trends of research topics in different eras [18]. The data in WoS show that scholars began to explore the digitization of ICH as early as 2004, while Chinese scholars began to speak out in the field in CNKI only two years later. From the annual dynamic trends in WoS, the number of international publications in the field of digitization of ICH was low and in a nascent state from 2004-2010; from 2011-2015, the number of literature grew year by year, especially rapidly from 2013-2015, and the research on digitization of ICH entered a period of development; from 2016-2019, the number of literature fluctuated and rose, reaching a peak in 2019, and followed by a slight decline in the next two years. In general, from 2016 to the present, research on digitization of ICH entered a deepening period [2].

From the Chinese core literature statistics of CNKI, the overall trend of Chinese research on digitization of ICH is basically consistent with the international trend, but the fluctuation of literature volume is more frequent and drastic, and the overall growth is more slowly. Although Chinese scholars began to publish the research results on ICH digitization in core journals only in 2006, the development was rapid, and a small peak occurred in 2009. But over the next

decade, global research on ICH digitization experienced significant growth, while the volume of literature in China grew slowly amidst fluctuations. After a plunge in 2018, CNKI significantly lags behind WoS in terms of the number of papers.

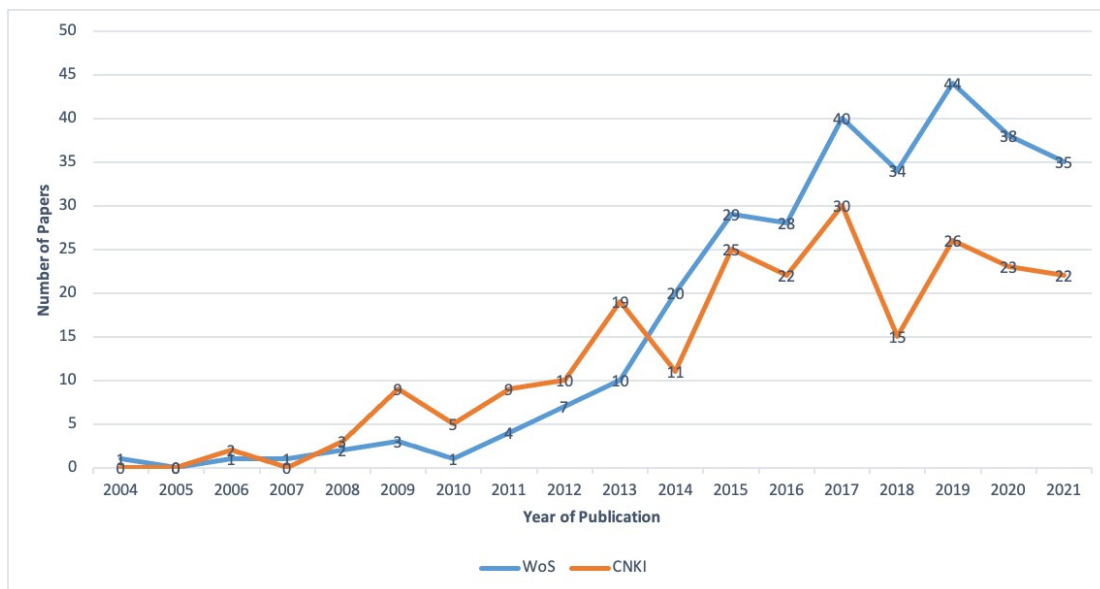


Figure 1. Yearly distribution of published papers on ICH digitization

3.2. Distribution of Disciplines

As an interdisciplinary discipline, the research on digitization of ICH involves a variety of disciplinary fields. Analyzing the disciplinary distribution of related research in China and abroad can show the participation of different disciplines in the research of digitization of ICH, and show the difference of knowledge structure in this field in China and abroad.

Both WoS and CNKI databases assign disciplinary categories to the papers. WoS contains a total of 252 disciplinary categories, of which 69 are involved in ICH digitization research; CNKI contains a total of 168 disciplinary categories, of which 27 are involved in ICH digitization research. Although the two databases have slightly different divisions of disciplinary categories, it is easy to find that both Chinese and the foreign studies on digitization of ICH cover many disciplinary categories, indicating that the field has attracted the attention of experts in various disciplines.

We consider the amount of papers in each discipline as a reflection of the participation of that discipline in the field of ICH digitization, thus defining the disciplinary participation rate calculation formula as follows:

$$R = \frac{N}{T} \times 100\%$$

In this formula, R represents the discipline participation rate; N means the number of papers published by a certain discipline; T means the total number of papers on ICH digitization. The top 20 disciplines in terms of participation rate according to the above formula in WoS and CNKI are respectively shown in Table 1 and Table 2.

As shown in Table 1, in the international context, Computer Science, including Computer Science Interdisciplinary Applications, Computer Science Theory Methods, Computer Science Information Systems, Computer Science Artificial Intelligence, Computer Science Software Engineering, Computer Science Cybernetics, occupies an important position in the research of ICH digitization, with a cumulative participation of 56.37%. And Table 2 shows that Chinese

computer science has published 38 papers in this field, with a participation rate of 15.83%. This contrasts sharply with the active performance of foreign computer science scholars, and Chinese scholars in the field of computer science should be more involved in the research of digitization of ICH in the future. WoS data show that many emerging engineering disciplines, such as Imaging Science Photographic Technology, Green Sustainable Science Technology, Environmental Sciences, Environmental Studies, Remote Sensing, Materials Science Multidisciplinary, etc. are increasingly involved in the digitization of ICH. While in China, the above-mentioned disciplines are still absent.

In China, research on the digitization of ICH is mainly concentrated in the cultural field, with a participation rate of 59.58%, while arts (including fine art, calligraphy, sculpture & photography, dramatic film & television arts, music & dance, with a cumulative participation rate of 17.91%) and sports (7.5% participation) also hold a high proportion. This may be due to China's rich cultural heritage, with a rich and specialized intangible cultural heritage in the arts and sports. In addition, social disciplines such as archives & museums, news & media, light industry & handicrafts, tourism also have a certain level of participation in the field of ICH digitization research in China. Although humanities multidisciplinary topped the list of all disciplines in WoS, the percentage of disciplines was only 17.79%, far less than that of China.

From the above comparative analysis of disciplinary distribution, it can be seen that in the field of ICH digitization research, engineering disciplines such as computer science and engineering electrical electronic disciplines are highly involved internationally, focusing on how to apply high technology to the field. In China, on the other hand, humanities and social science disciplines are more involved, focusing on the research of the digitization of specific ICH items.

Table 1. Top 20 disciplines of research on ICH digitization in WoS

Rank	Disciplines	No.	Participation Rate (%)
1	Humanities Multidisciplinary	53	17.79
2	Computer Science Interdisciplinary Applications	45	15.10
3	Computer Science Theory Methods	38	12.75
4	Computer Science Information Systems	33	11.07
5	Social Sciences Interdisciplinary	32	10.74
6	Information Science Library Science	29	9.73
7	Engineering Electrical Electronic	27	9.06
8	Computer Science Artificial Intelligence	23	7.72
9	Imaging Science Photographic Technology	19	6.38
10	Architecture	18	6.04
11	Computer Science Software Engineering	15	5.03
12	Green Sustainable Science Technology	15	5.03
13	Computer Science Cybernetics	14	4.70
14	Education Educational Research	14	4.70
15	Environmental Sciences	14	4.70
16	Archaeology	14	4.70
17	Art	13	4.36
18	Environmental Studies	12	4.03
19	Remote Sensing	11	3.69
20	Materials Science Multidisciplinary	9	3.02

Table 2. Top 20 disciplines of research on ICH digitization in CNKI

Rank	Disciplines	No.	Participation Rate (%)
1	Culture	143	59.58
2	Computer Software & Computer Applications	38	15.83
3	Archives & Museums	21	8.75
4	Library Information & Digital Library	20	8.33
5	Fine Art, Calligraphy, Sculpture & Photography	20	8.33
6	Sports	18	7.50
7	News & Media	14	5.83
8	Light Industry & Handicrafts	12	5.00
9	Dramatic Film & Television Arts	12	5.00
10	Music & Dance	11	4.58
11	Tourism	8	3.33
12	Literary Theory	5	2.08
13	Chinese Literature	4	1.67
14	Archeology	3	1.25
15	Internet Technology	3	1.25
16	Vocational Education	2	0.83
17	Publishing	2	0.83
18	Civil & Commercial Law	2	0.83
19	Automation Technology	2	0.83
20	Higher Education	2	0.83

3.3. Distribution of Countries

The literature on ICH digitization research in WoS come from 62 countries or regions, among which the top ten countries in terms of the number of publications are shown in Table 3. In terms of the number of papers, China ranks first with 87 papers, accounting for 29% of the total number of papers; followed by Italy, Greece, Spain, and the UK. All of the above countries have a long history and rich ICH, and thus have certain resource advantages in the research field of digitization of ICH.

However, CiteSpace analysis shows that China has a betweenness centrality value of 0, and Italy, Greece, and Spain also have low centrality values. The betweenness centrality represents the number of shortest paths through the node in a network. According to Freeman's (1977) definition, in the following formula, σ_{st} represents the number of shortest paths from node s to node t , and $\sigma_{st}(v)$ represents the number of nodes v passed in the shortest path from node s to node t . [19]

$$C_B(v) = \sum_{s \neq v \neq t \in V} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

In the knowledge map, the greater the betweenness centrality of a node, the greater the role it plays in the communication between other nodes [20]. The betweenness centrality of China is 0, which indicates the lack of international cooperation and communication in the digitization research of ICH in China.

CiteSpace's analysis data shows that the United States has the highest centrality value of 0.29, indicating that the United States plays an important role in the research of digitization of ICH. It is worth noting that Switzerland has a high centrality despite the small number of publications, indicating the country's outstanding contribution in the field. In general, the centrality values of all countries are not high, and three of the top ten countries in terms of

literature volume have a value of 0. Only the United States and Switzerland have a value greater than 0.2. This indicates that the exchange and cooperation among countries in the field of digitization of ICH is still low.

Table 3. Top 10 countries in ICH digitization

Rank	Country	Quantity	Centrality
1	PEOPLES R CHINA.	87	0
2	ITALY.	34	0.02
3	GREECE.	21	0.03
4	SPAIN.	19	0.07
5	ENGLAND.	15	0.13
6	USA.	12	0.29
7	SOUTH KOREA.	8	0
8	FRANCE.	8	0.04
9	TURKEY.	8	0
10	SWITZERLAND.	7	0.24

4. Analysis of Prolific Authors and Institutions

4.1. Analysis of Core Authors

The analysis of the authors of publications can clarify the distribution of core authors of ICH digitization research and help deepen the understanding of the disciplinary leaders of ICH digitization and the current status of disciplinary research [21].

According to the statistics, there are 781 authors of WoS literature, and the six highest posting authors are shown in Table 4. The highest publishing authors were Hui Zan (Cent China Normal Univ) and Kasper Rodil (Aalborg Univ), who published 5 papers. And there are 6 authors with 4 or more papers, 69 authors with 2 or more papers, and 712 authors who published only 1 paper, accounting for 91.2%. In CNKI, there are 357 authors, among which the highest author is Tan Guoxin (Cent China Normal Univ) with 7 papers. And there are 10 authors with 3 or more papers, 29 authors with 2 papers, and 318 authors with 1 paper, accounting for 89.9%. According to the data from WoS and CNKI, the majority of authors published only one paper. In accordance with the long-tail theory, most of the authors belong to sudden research and generally lack of continuity and depth of research.

The range of core authors in the field of ICH digitization can be determined according to Price's law, which is calculated as:

$$M \approx 0.749 \times \sqrt{N_{\max}}$$

In this formula, M is the number of articles published and N_{\max} is the number of articles published by the most productive authors. The authors with the number of articles reaching the M value can become core authors. If the number of papers published by core authors in the field reaches 50% of the total volume of papers, then a group of highly productive authors in the discipline has been formed [22].

Based on the above theory, the data of papers in WoS are analyzed, $N_{\max}=5$, and $M \approx 1.67$, which is rounded to 2. It is known that the core authors are those who have published 2 articles or more. According to the statistics, there are 69 core authors in WoS, and the total number of core authors in the field of ICH digitization only accounts for 19.2% of the total number of papers.

Using the same method to analyze the CNKI data, $N_{max}=7$, and $M \approx 1.98$, which is rounded to 2. The authors with 2 or more articles are considered core authors. The statistics show that there are 39 core authors. The total number of core authors only accounts for 30.2% of the total number of papers.

Both WoS and CNKI data show that the number of core authors does not reach 50%, which indicates that the field of ICH digitization is still in the development stage, and the authors are scattered and have not yet formed a stable core author group. However, the data from CNKI are significantly higher than those from WoS, indicating that the authors' publications are more concentrated in the field of digitization of ICH in China, and a higher proportion of scholars have conducted continuous and in-depth research. The international research on digitization of ICH has more scattered authors, and there may be more new forces joining the research in this field.

Table 4. List of top 6 prolific authors

Rank	Data in WoS			Data in CNKI		
	Author	Institution	No.	Author	Institution	No.
1	Hui Zan	Cent China Normal Univ.	5	Tan Guoxin	Central China Normal Univ.	7
2	Kasper Rodil	Aalborg Univ.	5	Zhang Ying	Shandong Univ.	5
3	Kate Hennessy	Simon Fraser Univ.	4	Tan Biyong	Shandong Univ.	5
4	Nikolaos Partarakis	Fdn Res & Technol Hellas	4	Huang Yonglin	Central China Normal Univ.	3
5	Fangtian Ying	Zhejiang Univ.	4	Zhou Yaolin	Wuhan Univ	3
6	Cheng Yao	Zhejiang Univ.	4	Feng Yun	Xizang Minzu Univ.	3

4.2. Analysis of Co-author Network

The analysis of co-author network can reflect the cooperative relationship among authors posting in the field of ICH digitization. Each node in the graph represents an author, the node size reflects the frequency of author co-occurrence, and the number and thickness of connecting lines represent the relationship and intensity of author cooperation [23].

In CiteSpace, the Node type is set to Author, the parameters, such as the time span, slice length, selection criteria etc., are as described in the study method, and the co-occurrence network map of authors in WoS is generated (see Fig. 2). There are 781 nodes 1563 connections in the map, and the network density is 0.0051. In the same way, we get the author co-occurrence network map in CNKI (see Fig. 3). There are 357 nodes, 258 links and 0.0041 network density in the map. The number of nodes, links and network density of the co-author network in WoS are significantly higher than those in CNKI, which shows that international scholars cooperate more frequently and closely than Chinese scholars in the field of ICH digitization.



Figure 2. Visualization of the co-author network in WoS



Figure 3. Visualization of the co-author network in CNKI
(Note: The authors names were translated for a better understanding.)

4.3. Analysis of Core Institutions

The institution network map was also generated by CiteSpace. Both in WoS and CNKI, the top ten institutions publishing literature are all from universities, and are dominated by comprehensive research universities (Table 5), which means the multidisciplinary intersectional environment of institutions has a positive impact on the research in this field. In terms of regional distribution, four of the top ten institutions in WoS are from China, three are from Greece, and one each from Italy, Denmark, and Canada. The top ten institutions in CNKI are located in various regions of China, among which Wuhan universities publish the most articles, and three of the four Chinese universities ranked in the top ten in WoS are from Wuhan. All three universities in Wuhan have set up research centers in the field of digitization of ICH, such as the National Research Center for Cultural Industries at Central China Normal University, the Research Center for Digital Cultural Heritage at Wuhan University, and Hubei University of Technology with the Hubei Research Center for Intangible Cultural Heritage, etc. These research platforms focus on digitization of intangible cultural heritage resources and provide strong support for the acquisition of research results.

Table 5. Distribution of core institutions in the field of ICH

Rank	Institutions in WoS	No.	Institutions in CNKI	No.
1	Zhejiang Univ (China)	7	Cent China Normal Univ (Wuhan)	16
2	Hubei Univ of Technol (China)	5	Shandong Univ (Jinan)	8
3	Cent China Normal Univ (China)	5	Tianjin Univ of Technol (Tianjin)	7
4	Politecn Milan (Italy)	5	Lanzhou Univ (Lanzhou)	6
5	Aalborg Univ (Denmark)	5	Sun Yat-sen Univ (Guangzhou)	6
6	Univ Patras Greece)	4	Renmin Univ of China (Beijing)	4
7	Univ Aegean (Greece)	4	Wuhan Univ (Wuhan)	4
8	Wuhan Univ (China)	4	Nanjing Univ (Nanjing)	4
9	Univ Crete (Greece)	4	Yunnan Univ (Kunming)	3
10	Simon Fraser Univ (Canada)	4	Zhejiang Univ (Hangzhou)	3

5. Keyword Visualization

Keywords are a precise summary of the ideas and contents of the literature, and their frequency, centrality and emergence and clustering are four important indicators for keyword analysis. Higher keyword frequency indicates higher attention, greater centrality indicates greater importance, clustering can explore the main directions of research in the field, and emergence can reflect the relatively prominent research hotspots in a certain time period [24].

5.1. Keyword Co-occurrence Analysis

The co-occurrence analysis of keywords in the literature on digitization of ICH can reveal the development trends and research hot topics in the field to a certain extent, which is conducive to grasping the focus in the development of digitization of ICH. In this study, the keyword co-occurrence mapping of 790 nodes, 1990 links, and 0.0064 network density was obtained for the documents screened in the WoS database, with the node type using keywords, a time span of 2 years, and TopN=50 (see Figure 4). Combining the list of high-frequency keywords (see Table 6), removing the subject terms related to the search strategy among them, and analyzing the co-occurrence frequency and centrality of keywords, it can be seen that the international hot keywords include virtual reality, augmented reality, digital preservation, digital humanities, design and digital technology, etc. The key co-occurrence mapping of 341 nodes, 422 links and 0.0072 network density was obtained for the keywords selected from the node types of the literature screened in the CNKI database, and it can be seen that the keywords in China include conservation, heritage, Tibet, communication, library, virtual reality, database and digital technology.

By comparing the similarities and differences of keyword concerns between the two types of sample literature, we can then understand the current status of research on the digitization of ICH in China and abroad. In terms of the overall number of network nodes and links, the distribution of research topics on digitization of ICH is more diverse and the keywords are more closely interconnected internationally, covering several hotspot areas. Among the top 10 high-frequency keywords, although WoS and CNKI both focus on the ways and effectiveness of digitization of ICH, there are differences in their research focus. Among the international studies, virtual reality and augmented reality appear more frequently and centrally, which shows that international research on ICH digitization is more focused on the study of digital technology. In China, research on ICH digitization is mainly focused on heritage and preservation, with the largest hotspot "protect", indicating that the most research is conducted for the purpose of ICH protection. In terms of centrality, "inheritance" is the most central, indicating that the perspective of ICH occupies a major position in domestic research.



Figure 4. Keyword co-occurrence mapping in WOS

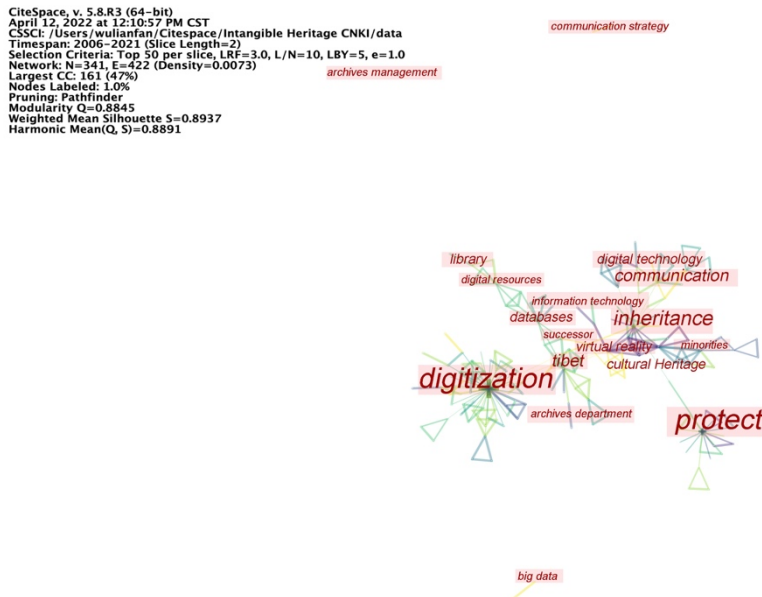


Figure 5. Keyword co-occurrence mapping in CNKI
 (Note: The keywords were translated for a better understanding.)

Table 6. Top 10 high-frequency keywords for ICH digitization research

Rank	WoS			CNKI		
	Keywords	Freq.	Centrality	Keywords	Freq.	Centrality
1	Intangible Cultural Heritage	75	0.59	Digitalization	42	0.25
2	Cultural Heritage	43	0.19	Protect	22	0.09
3	Digital Heritage	17	0.21	Inheritance	12	0.30
4	Intangible Heritage	17	0.12	Tiber	6	0.05
5	Virtual Reality	14	0.18	Communication	5	0.04
6	Augmented Reality	14	0.12	Library	4	0.01
7	Digital Preservation	8	0.06	Cultural Heritage	4	0.06
8	Digital Humanity	8	0.07	Virtual Reality	4	0.07
9	Design	7	0.07	Databases	4	0.06
10	Digital Technology	7	0.03	Digital Technology	4	0.01

5.2. Keyword Cluster Analysis

In this study, cluster analysis was performed on the basis of keyword co-occurrence analysis on the literature screened in WoS and CNKI databases, and keyword LLR cluster mapping was generated (see Figures 6 and 7). A total of 22 cluster labels were calculated from the cluster analysis of the sample literature on ICH digitization research WoS database. The Q value is 0.8848, which is much larger than 0.3 and has a reasonable structure; while in the Silhouette (S value), $S > 0.5$ means the clustering is more reasonable, and $S > 0.7$ means the clustering is convincing, and the S value of the top 10 clusters is higher than 0.7 as shown in Table 7. Among the top ten keyword clusters in WoS, #0, #2 are core search terms, #1, #3, #6, #7 can be categorized as research on the objects of digital preservation or development of ICH; #5, #9 can be categorized as research on digital collection and storage of ICH; #4 can be categorized as research on digital development technology of ICH, and #8 can be categorized as research on digital development theory of ICH.

The cluster analysis of the core papers of the CNKI database on the digitization of ICH formed a total of 88 cluster labels. The Q value is 0.8845, which is a reasonable structure, and the S value of the top 10 clusters is higher than 0.7. Among the top 10 keyword clusters in China, #0 is the

core search term, which is also an important way to protect ICH at this stage; #1 can be attributed to the research on the protection of ICH with ethnic characteristics; #2, #6, #9 can be attributed to the research on the implementation content of digital preservation of ICH; #3, #5, #8 can be attributed to the research on digital preservation of ICH and its strategies; #4, #7 can be attributed to the research on digital development of ICH.

By comparing the keyword clustering of the two types of literature, we can further understand the similarities and differences of the main research directions of digitization of ICH in China and abroad. In terms of the number of clusters, the clusters of Chinese ICH digitization research are more directional and richer, and although there is a certain degree of interlocking between the cluster labels, they are more scattered and not closely connected on the whole. In terms of Weighted Mean Silhouette, the international value is 0.99664 and the Chinese value makes 0.8937, the clustering of international studies on digitization of ICH is more clear and reasonable. In terms of the content of the clustering labels, international research on digitization of ICH differs from domestic research in that international research focuses more on specific digital technologies and has a more balanced research on technologies related to the preservation and exploitation of ICH, while in China the focus of research on digitization of ICH is more on the preservation level and relatively less on the consideration of exploitation.

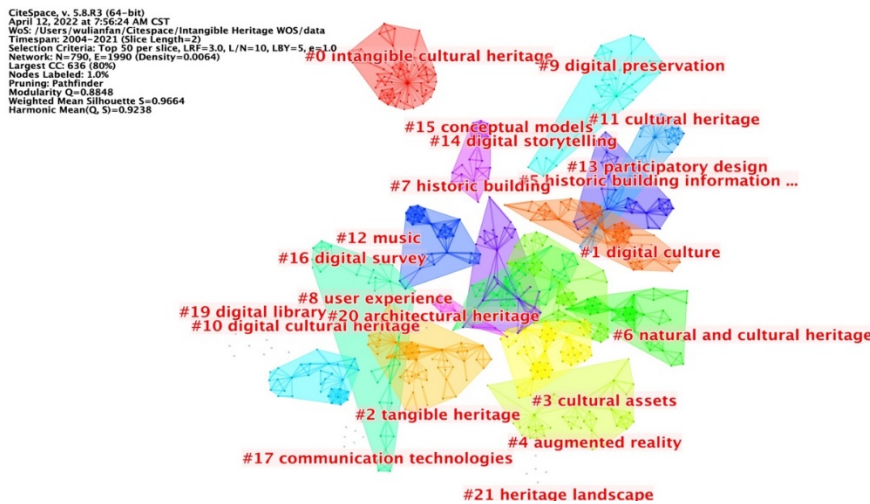


Figure 6. Knowledge map clusters in ICH digitization research in the world

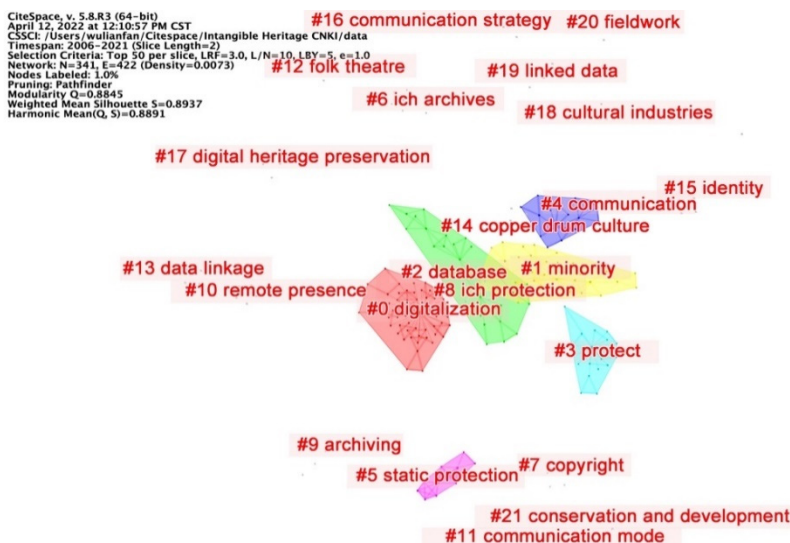


Figure 7. Knowledge map clusters in ICH digitization research in China
 (Note: The words were translated for a better understanding.)

Table 7. Keyword clusters in ICH digitization research

ID	WoS			CNKI		
	Cluster Label	Size	Silhouette	Cluster Label	Size	Silhouette
#0	Intangible cultural heritage	62	0.997	digitalization	53	1
#1	digital culture	52	0.97	minorities	32	0.985
#2	tangible heritage	51	0.972	databases	29	0.987
#3	cultural assets	50	0.94	protect	19	0.969
#4	augmented reality	39	0.941	communication	16	0.962
#5	historic building information modelling (hbim)	38	0.92	static protection	11	1
#6	natural and cultural heritage	37	1	ich archives	9	1
#7	historic building	36	0.918	copyright	6	1
#8	user experience	35	0.963	ich protection	5	0.972
#9	digital preservation	34	0.997	archiving	5	1

5.3. Citation Bursts Analysis

The Citation bursts refer to the keywords with a sudden increase in frequency within a certain period of time, and the analysis of the bursts can reflect the development direction and research frontier of ICH digitization. In the citation bursts analysis of the papers in WoS and CNKI databases, γ was set to 0.388 and 0.35, respectively, and the rest were selected as default parameters, and top 10 keywords with the strongest citation bursts were obtained, as shown in Table 8.

Table 8. Top 10 keywords with the strongest citation bursts in ICH digitization research

Rank	WoS			CNKI		
	Keywrds	Strength	Start& end year	Keywrds	Strength	Start& end year
1	Intangible Cultural Heritage	2.78	2012-2015	Cultural Heritage	1.46	2006-2013
2	Digital Archive	2.13	2014-2017	Virtual Reality	1.15	2006-2009
3	Digital Humanity	1.71	2014-2015	Information Technology	1.15	2006-2009
4	Smart City	1.53	2014-2019	Protect	2.77	2008-2015
5	Tangible Interaction	1.41	2014-2017	Archives	1.28	2010-2011
6	Tujia Brocade	2.37	2016-2019	Minorities	1.1	2010-2013
7	Historic Building	1.42	2016-2019	Manchu	1.04	2010-2013
8	Virtual Heritage	1.19	2016-2019	Utilisation	1.06	2014-2015
9	Participatory System	1.14	2016-2019	Communication Strategies	1.16	2018-2021
10	Information	1.38	2018-2021	Development	1.06	2018-2109

In the WoS database, in terms of burst intensity, words such as intangible cultural heritage, tujia brocade and digital archive have higher burst intensity, with 2.78, 2.37 and 2.13 respectively. In terms of burst time, the hotness of all burst words is maintained at 2-5 years, and there is no long-term existence of Research hotspots. Before 2012, the hotspots of scholars were scattered, and there were no obvious burst keywords, while after 2012, multiple burst keywords existed at the same time in different time periods, and there were multiple research hotspots; analyzing the specific burst keywords, we found that the word smart city (2014-2019) had a relatively long time span of 5 years, and the word information This indicates that the research on

capturing and representing ICH based on information computer technology has become an important frontier in the international research on digitization of ICH and has the value of further investigation.

In the CNKI database, in terms of burst intensity, only the word conservation has a high emergence intensity of 2.77, while the rest of the key words are below 1.5. From the perspective of emergence time, the heat of keywords are maintained at 2-8 in years, and the distribution of burst length is also relatively uniform; there existed several research hotspots before 2015, with 8 keywords, and after 2015, there are only 2 keywords and the intensity value is below the average, and the research heat is not high; the analysis of specific keywords reveals that cultural heritage (2006-2013), conservation (In addition, communication strategy and development are the two most recent keywords, and it can be assumed that the research on communication strategy and development of ICH is a recent research focus, in which the word communication strategy continues to this day and is a hot spot for future research in this field in China.

6. Conclusion and Outlook

In this study, we analyzed the papers in the field of ICH digitization in the WoS and CNKI databases with the help of CiteSpace tool, and analyzed the yearly distribution, the disciplines, countries, authors, institutions, and keywords, and the following conclusions were obtained:

(1) In terms of research stages and trends, the research on digitization of ICH can be roughly divided into three stages: the budding period (2004-2010), where the volume of papers are all small and the academic community begins to gradually pay attention to the field of digitization of ICH; the development period (2011-2015), where the volume of literature in WoS steadily grows and the volume of papers in China fluctuates; and the deepening period (2016-present), where the volume of papers oscillates up to a peak and then slowly declines. It is noteworthy that after entering the deepening period, the number of core papers in CNKI significantly lags behind the number of papers in WoS, and the gap is increasingly obvious. In this period, the research in the field of digitization of ICH has been perfected, and scholars are required to present newer ideas and more in-depth research results. WoS data show that engineering disciplines such as computer science, engineering electrical and electronics, and architecture occupy an important position in the research on digitization of ICH, and emerging disciplines such as green sustainable science technology, environmental sciences, environmental studies, remote sensing, and materials science multidisciplinary, are also applied in the field of digitization of ICH. However, the participation rate of computer science in CNKI papers is lower than the international level, and the aforementioned emerging engineering disciplines have not yet been involved in the research in China. This may be the reason for the increasing gap between the number of publications on ICH digitization research in China and the international ones in recent years. From a global perspective, an important feature of modern knowledge production is the development trend of synthesis and socialization based on a high degree of differentiation and extensive penetration, and the crossover, integration and interpenetration of knowledge are occurring widely in different disciplines. In the future, the research in the field of digitization of ICH in China should follow the international development trend, strengthen the cross-fertilization between different disciplines, encourage more engineering disciplines to participate in the research, and focus on the research and application of emerging technologies.

(2) In terms of cooperation and exchange, there is less cooperation among countries in the field of ICH digitization research. Among the top ten countries in terms of literature volume, only the United States and Switzerland have centrality values greater than 0.2, and three countries have centrality values of 0, including China. China is the most published country in the field of ICH digitization, accounting for 29% of the total volume of papers, but its low centrality value

indicates that Chinese research is more independent and lacks international cooperation and exchange. This may be due to two reasons: first, Chinese culture is vast and profound, and the field of ICH involves ethnic and regional expertise, which makes it difficult for foreign experts to conduct in-depth research; second, in terms of disciplinary distribution, Chinese research focuses on the digitization of specific ICH items, while the international research trend is to apply high technology to ICH items. The two differ in the focus of their research, and the Chinese research direction has more limited appeal and reference to international colleagues. The study also found that the lack of collaboration is not limited to domestic and foreign scholars, but also among local scholars, through the analysis of author co-occurrence using CiteSpace. In contrast, international collaboration among authors in the field of ICH digitization is more frequent and closet. In the future, China should stimulate the communication vitality of various research subjects, build an open and integrated innovation ecology, and jointly promote the construction and deepening of a collaborative network for research on the digitization of ICH.

(3) In terms of core authors, there is no stable core group of authors in the field of ICH digitization in either China or foreign countries. Chinese authors are more concentrated in their publications, and a higher proportion of scholars have conducted continuous and in-depth research. In the international research on ICH digitalization, the authors of the articles are relatively scattered, and there are more new forces joining the research in this field.

(4) In terms of research institutions, the data of WoS and CNKI both show that comprehensive colleges and universities occupy a dominant position in the research on the digitization of ICH. And some science & engineering colleges and normal colleges also have relatively prominent performance. This shows that the multidisciplinary environment is conducive to the in-depth research on the digitization of ICH. Some Chinese normal universities (Cent China Normal Univ) and science & engineering colleges (Hubei Univ of Technol) are outstanding in the field of intangible cultural heritage digitization. These institutions have relatively mature scientific research platforms in the field of intangible cultural heritage, which provides strong support for the output of research results in this field. In the future, China should continue to support the establishment and development of more ICH research platforms and form more ICH research teams with more diverse backgrounds and more reasonable echelons.

(5) In terms of keywords, there are 790 keywords in the sample literature of WoS, and their fields of concern are relatively rich and their perspectives are diverse. In terms of high-frequency keywords, although its content covers a wide range, it mainly focuses on the fields of computer and engineering technology. After further sorting out, it is found that the international research path of ICH digitization mainly relies on in-depth research on related disciplines, and then explores and forms its own research theory and practice system. Its research themes are relatively concentrated and have strong correlation and continuity with each other, and a relatively complete discipline system has been initially formed. The number of keywords generated by the sample literature in CNKI is only 341, which is not as deep and broad as international research. In terms of high-frequency keywords, Chinese literature is more uniform in the selection of topics, and tends to carry out research on humanities and computers at the same time. In addition, China's research on digitization of ICH is more likely to be guided by national and local policies, so there is a phenomenon of blindly following the trend in the choice of research topics. It has seriously affected the improvement of the Chinese research system in the field of ICH digitalization. In the future, scholars of ICH digitization need to strengthen the research and reference of relevant foreign basic theories, deeply cultivate, deepen the understanding of the regularity of digitization of ICH, and jointly build a complete and solid basic research system for ICH digitization.

7. Conclusion

With the rapid development of information technology, the digitization of ICH has ushered in an opportunity for rapid development. This research uses CiteSpace software to visually analyze and summarize the international and Chinese ICH digitization papers, and further sort out and reveal the research progress, hotspots and trends of ICH digitization. This will guide the direction of the follow-up research and practice work on the digitization of ICH in China by identifying and analyzing the research status of ICH in China and abroad, and then promoting the protection, dissemination and inheritance of ICH.

There are still some limitations in this study. First of all, the research selects the method of subject keyword retrieval to collect relevant papers of ICH digitization, which is difficult to avoid the occurrence of omissions in a small number of papers retrieval. Secondly, because CiteSpace software pays more attention to the analysis of numbers, it is easy to cause a few newly published important results to be ignored by the software due to insufficient attention during analysis. The above problems will affect the representativeness and comprehensiveness of the research results to a certain extent. In the future, on the basis of analyzing the results with the help of CiteSpace software, the research will consult more experts in related fields to further confirm and supplement the results of this research.

Acknowledgments

This research was supported by 2021 Wenzhou Basic Scientific Research Projects (Self-funded) [Project NO.: R20210047] and "14th Five-Year Plan" Teaching Construction and Teaching Reform Research Project of Wenzhou Polytechnic [Project NO.: WZYZD202101].

References

- [1] Zhang Q. and Wang D., "Intellectual structure and research frontier of international intangible cultural heritage studies", *Jiangsu Science & Technology Information*, vol.36(2019), no.32, p.12-16.
- [2] Shen Y., "Study on the protection of intangible cultural heritage abroad and its inspirations", *Journal of Hunan University of Humanities, Science and Thechnology*, vol.38(2021), no.01, p.62-66.
- [3] China's Intangible Cultural Heritage List Project [DB/OL]. [2021-12-19] [http:// www. ihchina. cn/ chinadirectory.html#target1](http://www.ihchina.cn/chinadirectory.html#target1).
- [4] LI X. and SUN Z., "Research on the knowledge map and visualization of fashion design field in China based on CiteSpace", *Journal of Silk*, vol.57(2020), no.05, p.25-34.
- [5] Liu H. and Liu Y., "Research on Chinese intangible cultural heritage archives based on CiteSpace", *Public Communication of Science & Technology*, vol.12(2020), no.20, p.36-38.
- [6] Chen C. and Huang Y., "Knowledge map analysis of domestic intangible cultural heritage research based on CiteSpace", *Packaging Engineering*, vol. 41(2020), no.14, p.228-234.
- [7] Duan X., "Bibliometric analysis of CNKI intangible cultural heritage research from 2001 to 2020", *Cultural Heritage*, 2021, no.04, p.28-36.
- [8] Li Y., Dai G., and Zhang J., "Recent ten years (2007–2017) research hotspots, trends and frontiers of Chinese Wushu culture: visual analysis based on knowledge map", *Journal of Nanjing Sports Institute*, vol.1(2018), no.10, p.69-80.
- [9] Chen H., "CiteSpace analysis of Chinese Wushu research in the past twenty years", *Taiyuan University of Technology*, 2019.
- [10] Xing W., "A summary of the research on sports intangible cultural heritage based on CiteSpace III". *Wushu Studies*, vol.5(2020), no.09, p.107-110.
- [11] Liu M. and Zhang Y., "Bibliometric analysis of sports intangible cultural heritage researches in china based on knowledge maps", *Hubei Sports Science*, vol.40(2021), no.05, p.382-388+413.

- [12] Huang J. and Wang J., "Bibliometric and visual analysis of researches on Chinese traditional Wushu in the past 20 years", *Wushu Studies*, vol.7(2022), no.02, p.61-65.
- [13] Huang M., "The academic layout and prospect of the research on the protection of minority traditional culture (2000-2019)", *Journal of Guangxi Institute of Socialism*, vol.31(2020), no.03, p.65-74.
- [14] Li Y. and Wu Y., "The development and trends of research on Tibetan Opera in the past three decades: based on a quantitative visual analysis", *Journal of Ethnology*, vol.12(2021), no.04, p.101-109+122.
- [15] Wang J., "Atlas analysis of the data of the research on the intangible cultural heritage of ethnic minorities in China during 2015-2021: A case study based on 1144 articles from CNKI", *Journal of Hezhou University*, vol. 37(2021), no.04, p.7-14.
- [16] Zhang, D., et al. "Study on sustainable urbanization literature based on Web of Science, scopus, and China national knowledge infrastructure: A scientometric analysis in CiteSpace." *Journal of Cleaner Production* 264, 2020.
- [17] Chen, C. "Searching for intellectual turning points: Progressive knowledge domain visualization", *Proceedings of the National Academy of Sciences*, vol.101(2004), suppl. 1, p.5303-5310.
- [18] Bai Y., "Aimaitijiang·Abuduhali, Deng Feng. Analysis of scientific knowledge map of tourism transportation research in china", *Enterprise Economy*, no.12, p.136 -142, 2018.
- [19] Brandes U. "A faster algorithm for betweenness centrality", *Journal of Mathematical Sociology*, vol.25(2001), no.2, p.163-177.
- [20] Su X., Li X. and Kang Y., "A Bibliometric analysis of research on intangible cultural heritage using CiteSpace", *SAGE Open*, 9, 2019.
- [21] Zeng X., LU D. and He B., "Review of research on china's tourism sharing economy--based on CiteSpace visual literature analysis", *Resource Development & Market*, vol.34(2018), no.8, p.1179-1184.
- [22] Xu F., "Analysis on the research status of art field in digital context: based on cnki data (1998 - 2019)", *Journal of Southwest Minzu University (Humanities and Social Science)*, vol.42(2021), no.04, p.234-240.
- [23] Garfield E., "In tribute to Derek John de Solla Price: a citation analysis of little science, big science", *Scientometrics*, vol.7(1985), no.3, p.487-503.
- [24] CHEN Y., CHEN C., and Hu Z., et al, "Principles and applications of analyzing a citation space", Beijing: Science Press, 2014.