

Progress and Trends in Digital Literacy Research - Based on Knowledge Map Analysis Using Cite Space and VOSviewer

Zhi Sheng

Viterbi School of Engineering University of Southern California, California, USA

Abstract: Information and digitization have increasingly become dominant in human existence and practice. The advent of the digital age brings opportunities and challenges for a digital transformation in global education, facing a trend characterized by digitalization, intelligence, and wisdom. Systematically reviewing the current state of digital literacy research and exploring core institutions, research ‘hotspots’, and cutting-edge studies are necessary and meaningful. Knowledge map analysis methods based on knowledge metrics, utilizing tools like Cite Space and VOSviewer, offer effective solutions. This involves quantitative analysis of digital literacy literature from the Web of Science and CNKI databases, exploring characteristics of journal publications, examining influential researchers, mapping international collaboration among countries, and identifying research hotspots and frontiers by analyzing keyword co-occurrence.

Keywords: Digitization; Digital Literacy; Cite Space; VOSviewer

1. Introduction

Digital literacy, an evolution and expansion of concepts like media literacy and information literacy in the digital age, is a vital skill for modern citizens to fully participate in the digital society. Currently, the burgeoning research in digital literacy is making it a rapidly evolving knowledge field. China’s research in digital literacy started later and is relatively limited in volume. While domestic attention to digital literacy has increased in recent years, research remains inadequate, often focusing on the progress of digital literacy education in certain countries or analyzing specific digital literacy frameworks, lacking a comprehensive understanding of the field. This paper aims to objectively address questions like which authors and institutions are dedicated to digital literacy research, what the research hotspots are, and how digital literacy relates to other literacy concepts. By clarifying the current state, knowledge base, and research themes of international digital literacy research, this study provides valuable insights for future research. The paper conducts a scientometric analysis of literature on digital literacy since 2013, integrating and organizing its distribution over time, space, core literature, and journals, summarizing the current state and hotspots of digital literacy research, and offering a reference for future digital literacy research in China.

2. Data Sources and Research Methods

The Web of Science is widely acknowledged as the world’s most reliable and renowned citation index (Olawumi & Chan, 2018). It’s considered more prestigious than other databases due to its comprehensive and reliable information and accurate evaluation metrics for journals, authors, and articles (Gholampour et al., 2022). To ensure the representativeness and academic nature of literature data, this study selected the Web of Science Core Collection and its three effective indexes: Science Citation Index Expanded (SCIE), Social Sciences Citation Index (SSCI), and Arts & Humanities Citation Index (AHCI), to review current literature on digital literacy. Search terms “digital literacy” and “digital competence” were used with labels (TI, KY), Boolean operators (OR, AND), exact phrase terms (“), and wildcard (*) symbols for the period 2013–2023, filtering out unrelated articles to obtain 9382 documents as the research sample.

Simultaneously, in CNKI (China National Knowledge Infrastructure), using “Subject=Digital Literacy” as the search criterion and the timeframe of 2013–2023, irrelevant documents like brief reports and calls for papers were excluded. This resulted in 5404 documents related to digital literacy, with 668 from the “Chinese Social Sciences Citation Index (CSSCI)” database. These documents

were exported in Refworks format. To meet the data format requirements for creating knowledge maps, the data was converted using Cite Space's data conversion function, preparing it for subsequent analysis.

3. Knowledge Landscape of Digital Literacy Research

3.1 Distribution Characteristics

3.1.1 Annual Publication Volume

The annual distribution of literature on digital literacy in both Chinese and English, as illustrated in Figure 1, shows an upward trend. The volume of English publications significantly exceeds that of Chinese ones, especially after 2017, with a marked increase in English publications starting from that year, while a significant increase in Chinese publications began in 2021.

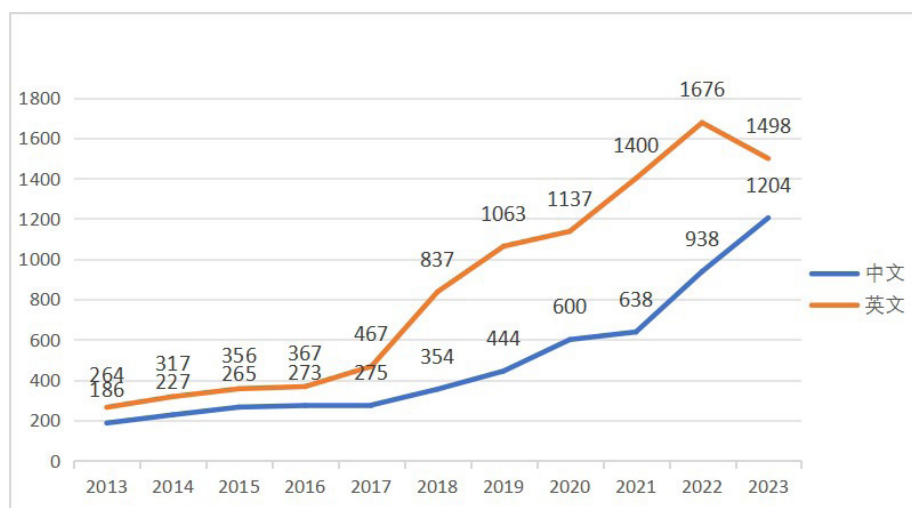


Figure 1. Annual Publication Volume of Digital Literacy Literature in Chinese and English

3.1.2 Distribution of Research Authors

By running Cite Space software and selecting the Author option for analysis, the study identified authors with a high volume of publications within the set period in the field of digital literacy. The visualization analysis results reveal academic collaborations among digital literacy scholars outside China, as shown in Figure 2. However, most domestic scholars researching digital literacy have not formed a collaborative network centered on this field, indicating the absence of a concentrated core group of research authors in digital literacy in China.

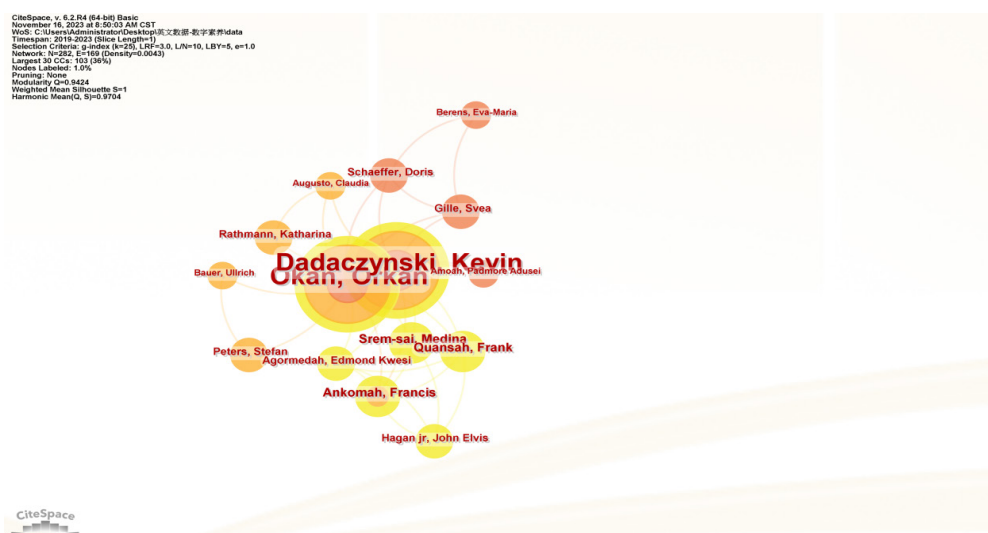


Figure 2. Collaboration Among International Research Authors

According to Price's Law, which states that authors with three or more publications are considered prolific, the results from Cite Space reveal that from 2013 to 2023, there were 17 prolific authors in China, with the highest number of publications being only 7. In contrast, there were 58 prolific international authors, with the highest number of publications reaching 23. For more details, see Table 1.

Table 1. Prolific Authors and Their Publication Volumes Domestically and Internationally

No.	Chinese Authors	Number of published papers	Foreign Authors	Number of published papers
1	Di Wu	7	Dadaczynski, Kevin	23
2	Yueliang Zeng	7	Okan, Orkan	23
3	Youqun Ren	6	Kucirkova, Natalia	13
4	Zhang Xiong	6	Aguaded, Ignacio	12
5	Yan Liu	6	Tomczyk, Lukasz	11
6	Ruhua Huang	6	Cabero-almenara, Julio	11
7	Sha Zhu	5	Jiang, Lianjiang	8
8	Xiaozhe Yang	5	Osborne, Richard H	6
9	Junhong Xiao	4	Li, Mimi	6
10	Guoshuai Lan	4	Garcia, Antero	6
11	Ning Sun	4	Cheng, Christina	6
12	Ping Wan	3	Bacalja, Alexander	6
13	Ping Ke	3	McGrew, Sarah	5
14	Wenyun Liu	3	Chu, Samuel Kai Wah	5
15	Youmei Wang	3	Chaudhary, Kaylash	5
16	Xiaolong Qian	3	Altschwager, Darcey	4
17	Shaoqing Guo	3	Amgott, Natalie	4

3.1.3 National Distribution

To better understand the countries/regions with high publication rates, data was imported into VOS viewer with the country as the node type. In the visualization results, larger nodes indicate more publications related to digital literacy in that country/region. Figure 3 clearly shows that the four countries/regions with the highest number of publications are the United States, United Kingdom, Australia, and Spain. The status of scientific cooperation networks in digital literacy research across countries, based on the organizational affiliations of authors, indicates over 30 countries collaborated scientifically in publishing. These collaborations led to 1504 co-authorships and 1723 publications in the field. Overall, the collaboration results reveal significant scientific cooperation in the United States (57 links), United Kingdom (41 links), Spain (38 links), Australia (33 links), Canada (33 links), and Italy (21 links). The most substantial collaborations were between the USA-UK, USA-Australia, USA-Greece, and USA-Spain. China's collaborations with other countries in both number and intensity are continually increasing.

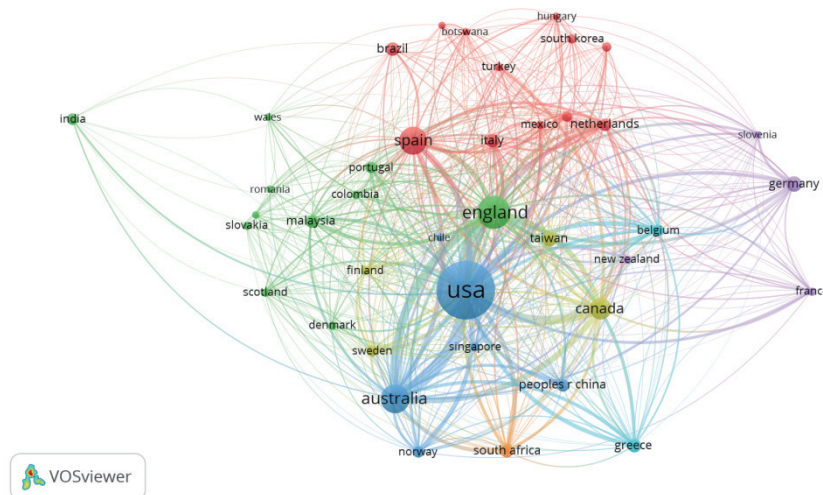


Figure 3. Distribution of the countries that issued the document

3.2 Co-occurrence Map Analysis

3.2.1 Keyword Frequency Analysis

Keywords represent the core essence of literature, and analyzing their frequency can help understand the research hotspots and focal points of a field. In the map, a greater number of connecting lines indicates stronger relationships between keywords. As shown in Figure 4, the top ten keywords derived from digital literacy include: digital economy, digital divide, digital villages, information literacy, rural revitalization, digital technology, libraries, artificial intelligence. These are the key focuses and hotspots in this area of research.

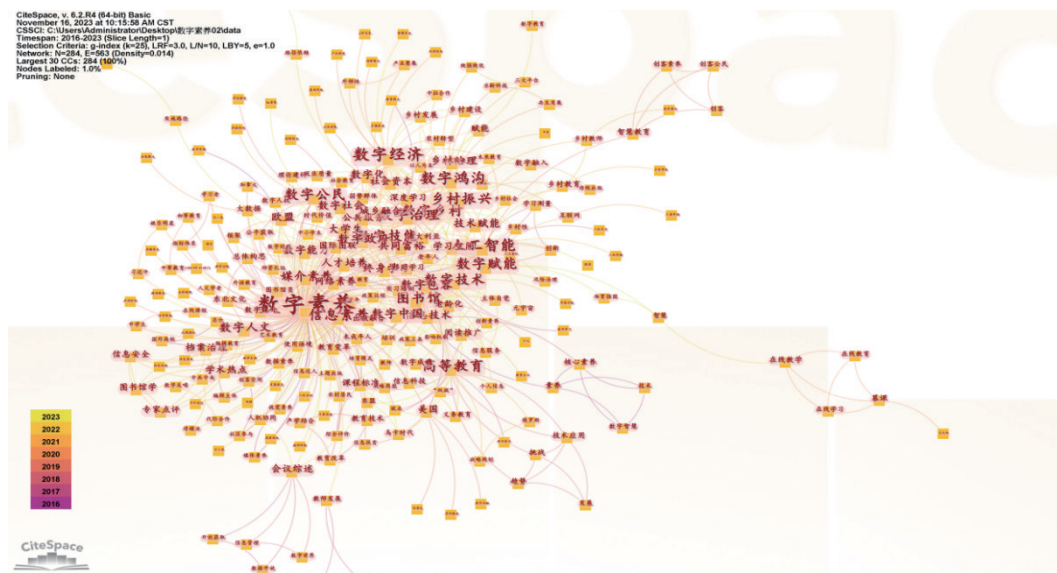


Figure 4. CNKI high-frequency keyword map

The VOSviewer bibliometric tool uniquely presents various network views like clustering, tagging, and density, ensuring aesthetic visuals while maintaining accuracy in the co-occurrence relationships between nodes. In the process of drawing the co-occurrence knowledge map for English keywords, 9,382 foreign literature pieces contained 31,544 keywords. Figure 5, created by setting the minimum visibility for keywords to 8, reveals that besides search terms, keywords such as “digital literacy,” “digital competence,” “media literacies,” “teacher education,” and “university students” are frequent, highlighting the international focus on digital competencies and media literacy research. The figure also indicates significant research attention on teachers, students, and government roles in enhancing digital literacy.

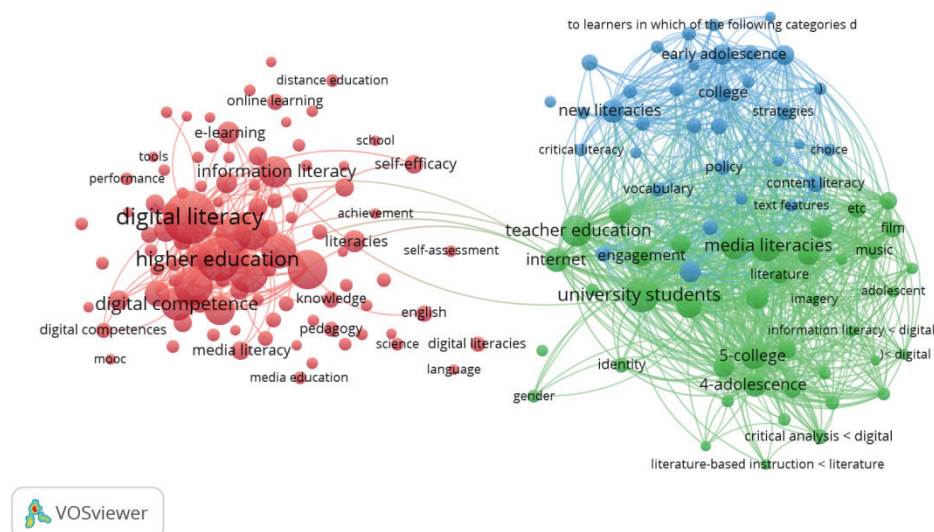


Figure 5. Web of Science high-frequency keyword map

3.2.2 Keyword Clustering Analysis

Keyword clustering analysis simplifies the co-occurrence network of keywords into fewer clusters using cluster statistical methods. A Modularity Q value (Q value) over 0.3 indicates significant clustering. In the CNKI keyword clustering map, “Modularity Q=0.596” signifies significant clustering in digital literacy. “Silhouette S=0.8596” suggests the results are reliable. Figure 6 shows 12 clusters, including #0 Digital Literacy, #1 Digital Villages, #2 Digital Divide, #3 Digital Citizenship, #4 Artificial Intelligence, #5 Libraries, #6

Higher Education, #7 Curriculum Standards, #8 Digital Humanities, #9 Conference Reviews, #10 United States, and #11 MOOCs.

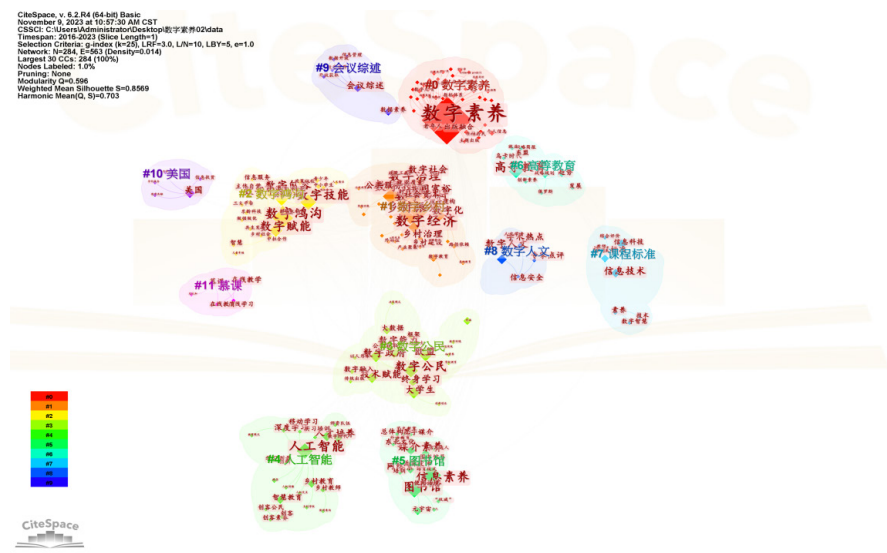


Figure 6. CNKI keyword clustering diagram

In the Web of Science keyword clustering map, “Modularity Q=0.5313” and “Silhouette S=0.8096” also indicate the reliability of the clustering results. The map displays 6 clusters: #0 Digital Competence, #1 Social Studies, #2 Digital Literacy, #3 Digital Tool, #4 Technology-Induced Stress, #5 Pedagogical Affordance, and #6 Language Assessment Literacy.



Figure 7. Keyword clustering map of Web of Science

3.2.3 Keyword Timeline Analysis

The keyword timeline analysis arranges keywords in chronological order to show their evolution. Each cluster corresponds to a line, with nodes on the line representing the main research content covered by that cluster, as shown in Figure 8. Comparisons reveal that different clusters have distinct main contents and durations. Research on digital literacy has been consistent throughout, with related areas like digital divide, digital governance, and digital empowerment emerging as recent hotspots.

4. Research hot topics

4.1 Digital Literacy Conception

Digital literacy, as defined by Gilster in 1997, is the ability to understand and use the vast and complex information available through computers. Evolving with technological advancements, it now encompasses elements from information and communication technology literacy, information literacy, and media literacy. The American Library Association (2013) described it as the capacity to use ICT to find, understand, evaluate, create, and communicate digital information, requiring both cognitive and technical skills. JISC (2014) offered a broader European perspective, defining it as the competencies suitable for living, learning, and working in a digital society, encompassing technical, socio-technical, cognitive, and behavioral aspects. However, these definitions miss the complexity of

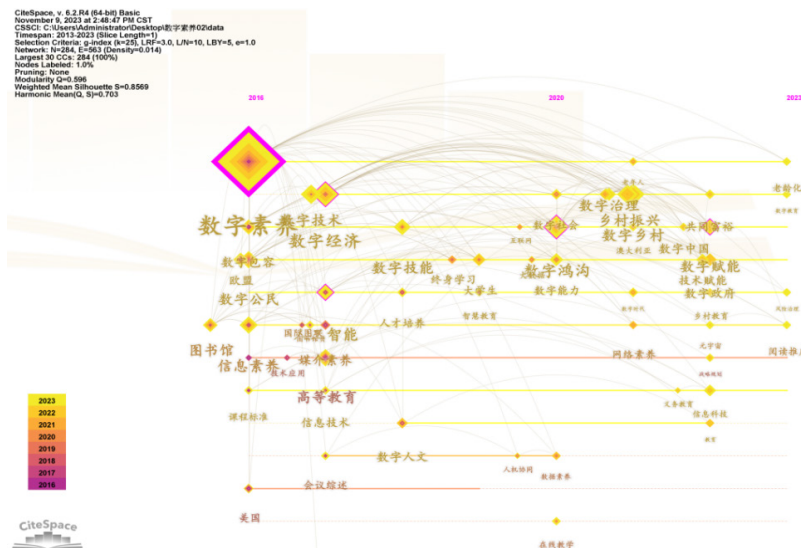


Figure 8. Keyword timing map

digital literacy within specific disciplines or professions and elements like critical thinking, creativity, and confidence, which Belshaw (2012) considers fundamental to digital literacy.

4.2 Digital Divide

The digital divide is a metaphorical term with uncertain and open-ended connotations. Originating from a U.S. report by the NTIA (1999), it was characterized as a significant economic and civil rights issue. Arquette (2001) described it as disparities in personal computer ownership across social groups during the Apple II era. The digital divide encompasses gaps in ICT infrastructure, ownership, and usage. Van Dijk and Hacker (2003) noted that the divide suggests two distinct groups and cautioned that most ICT inequalities are relative, not absolute. Initially, the digital divide described societal differences arising from digital media use. Chinese research initially focused on information dissemination and library science, defining the divide as differences in internet usage across social strata. As the digital divide's potential impact on social equity became more evident, legal studies emerged, leading to policy responses in China. The 13th Five-Year National Informatization Plan (2016) included the digital divide as a key issue, highlighting it as a significant shortfall in China's informatization progress and making its reduction a goal for "Digital China" development. The concept of the digital divide is fluid, evolving with technological advancements, and there's a need to be vigilant about new forms of divides emerging from new technologies.

In summary, the digital divide, or "Digital divide," refers to the information gap and increasing polarization of wealth caused by disparities in information access, technology application, and innovation capability during the global digitalization process. It exists both between and within countries and regions. Bridging the digital divide in the digital era is crucial not only for helping disadvantaged groups keep pace with the times but also for implementing national development strategies. This requires enhancing information infrastructure and improving digital literacy across the population.

4.3 Digital Governance

The rapid development of science and technology has reduced learning barriers and provided auxiliary tools. The application of modern technological innovation has brought new opportunities to the field of education. With the intervention of information technology, behavior research based on multiple data platforms is gradually emerging from the fields of education and research (Hefnan, 2011). In recent years, utilizing educational data mining and learning analysis to improve the level of educational governance has become a hot topic in this field (Zhang et al. 2018). For example, there are three main methods for evaluating the learning process of students: Educational Data Mining (EDM), Machine Learning, and Statistical Analysis (Draxler and Greleler, 2012; Petkovic, 2016). Based on specific data analysis, these methods can be divided into rule mining and fuzzy representation (Angeli et al. 2017), and can also be divided in other ways based on different foundations (Agudo Peregrina et al. 2014). A highly cited study used online data from Learning Management Systems (LMS) to predict student behavior and intervene in teaching (Macfadden and Dawson, 2010), which includes not only content data but also personality data.

4.4 Digital Empowerment

Digital empowerment refers to empowering enterprises, organizations, or individuals with more abilities and opportunities through digital technology and means to achieve more efficient, innovative, and intelligent work and lifestyle. Digital empowerment

plays an increasingly important role in promoting economic development, digital transformation of enterprises, changing the industrial landscape, and restructuring government governance. Digital empowerment has had a huge impact on people's work and life, involving numerous fields such as agriculture, commerce, education, healthcare, media, and so on. Thomas (1990) studied the cognitive elements of digital empowerment and proposed that the cognitive model of digital empowerment includes four elements: influence, ability, meaning, and selectivity. Makinen (2006) pointed out that digital empowerment refers to empowering specific populations through digital tools such as big data, mobile internet, and artificial intelligence, enabling them to acquire corresponding life skills and survival abilities. From a sociological research perspective, Hermansson E and M Sartensson L (2011) point out that digital empowerment is the ability to identify, promote, and enhance people's ability to respond to needs and solve their own problems, and mobilize the necessary digital resources to make people consciously control their lives. Chinese scholar Li Xiaoyu (2018) believes that digital empowerment refers to empowering people through digital technologies such as cloud computing, big data, the Internet of Things, mobile internet, and artificial intelligence. Digital empowerment promotes rural revitalization and promotes high-quality rural development (Qin Qiuxia, 2021). Digital empowerment has achieved innovation in government governance tools and improved governance efficiency (Zhao Miao, 2022). Hu Weiwei and Shen Wenjing (2022) believe that the development of digital technology has intensified the emergence of the concept of "digital empowerment". Digital empowerment mainly refers to using digital technology means to provide diverse service methods for action subjects, enhance their service capabilities, tap into their service potential and autonomy, and ultimately effectively achieve established goals.

5. Conclusion

This study used bibliometric methods to analyze 9382 digital literacy themed papers included in the core collection of ScienceNet from 2013 to 2023, in addition to 668 Chinese core journal papers. Based on this, a visual map of digital literacy research was drawn, examining the evolution of basic characteristics, research hotspots, and trends in this field. The results indicate the following points:

- *The number of publications in the literature on digital literacy research is showing a fluctuating upward trend, and it is expected that the number of publications in this field will continue to increase in the coming years.

- *The United States is a world leader in digital literacy research.

- *Although the core group of authors has not yet formed, some authors, such as Dadaczynski, Kevin, and Okan, have already had influence in this field. The main research hotspots are the connotation of digital literacy, digital divide, digital governance, digital empowerment, etc.

Based on the above findings, this study suggests that future research on digital literacy can focus on the following three areas: firstly, existing relevant literature mainly focuses on the study of the teacher-student group. Therefore, studying external factors that affect the implementation of digital literacy, such as policies, environment, and culture, can be strengthened. Secondly, some scholars have established relatively fixed research topics and should strengthen the construction of diverse methods and paradigms, enhance international and regional cooperation, and promote research innovation in topics, methods, and paradigms. Thirdly, the integration of theory and practice can be strengthened, and research should be oriented towards practical needs to promote the formal effectiveness of digital literacy in practical classrooms.

References:

- [1] Coldwell-Neilson J .Digital literacy—a driver for curriculum transformation[C]//Research & Development in Higher Education: Curriculum Transformation.2017.
- [2] Hermansson E, Metensson L.Empowerment in the midwifery context---Aconcept analysis[J].Midwifery, 2011, 27 (6) :811-816.
- [3] Hu, W. & Shen, W. The Practical Logic and Operational Mechanism of Technology Empowering Rural Digital Governance: An Empirical Study on the Construction of Digital Countryside in Guanzhong H Village [J]. Journal of Hunan Agricultural University (Social Science Edition), 2022, 23 (05): 61-67+75
- [4] Makinen M.Digital empowerment as a process for enhancingcitizens' participation[J].E-learning, 2006, 3 (3) :381-395.
- [5] Qin, Q., Guo, H. & Zeng, Y. Digital Empowerment and Implementation Approaches in Rural Revitalization [J]. Journal of Jiangsu University (Social Science Edition), 2021,23 (05): 22-33
- [6] Thomas K W, Velthouse B A.Cognitive elements of empowerment:An“interpretive”model of intrinsic task motivation[J].Academy of Management Review, 1990, 15 (4) :666-681.
- [7] VAN DIJK J,HACKER K.The digital divide as a complex and dynamic phenomenon[J].The information society,2003(4):315-326.
- [8] Zhao, M., Bao, J. & Liu, Y. From Empowerment to Inclusion: Mechanisms for Generating and Resolving the Non equilibrium Dilemma in Digital Government Construction [J]. China Administrative Management, 2022 (12)