



Evolution and diffusion of information literacy topics

Yating Li¹ · Ye Chen² · Qiyu Wang¹

Received: 15 September 2020 / Accepted: 24 February 2021 / Published online: 12 March 2021
© Akadémiai Kiadó, Budapest, Hungary 2021

Abstract

Investigation of the topic of information literacy and its changes can be informative for researchers and provide a better understanding of the corresponding domains. This study conducted a topic model dynamic analysis of the articles on information literacy studies in the Web of Science core collection database that were published from 2005 to 2019. The global topics and their popularities, topical similarities and correlations, along with the evolution of temporal topics and the diffusion of subject local topics were analyzed and presented. Nine global topics differed in terms of their temporal and subject characteristics, and this study focused on ability, technology, field, people, place and application of information literacy. For the temporal local topics, crossing was the main evolutionary mechanism; hence, the core topic words were relatively stable, but few new research directions have been explored in recent years. For the subject local topics, absorbing with division and absorbing were the main mechanisms, which supported the diffusion progress of information literacy studies among subjects. However, it is necessary to promote the development of future research through the innovative development of multidisciplinary integration. Researchers and practitioners should focus on the impact of information technology, increase the breadth and depth of the research field, and develop innovative evaluation methods that are based on data to promote the comprehensive, sustainable and effective improvement in information literacy.

Keywords Information literacy · Topic modeling · Topic evolution and diffusion · Word embedding

Introduction

The President of the American Information Industry Association, Paul Zurkowski, in 1974, said, “Information literacy is a kind of skill that comprehensively uses information tools and information sources to solve problems encountered” (Zurkowski, 1974). In 1989, the American Library Association Presidential Committee noted, “To be information literate, a

✉ Ye Chen
cheny@ccnu.edu.cn

¹ National Engineering Laboratory for Educational Big Data, Central China Normal University, Wuhan, China

² School of Information Management, Central China Normal University, Wuhan, China

person must be able to recognize when information is needed and have the ability to locate, evaluate, and use effectively the needed information” (ALA, 1989). Since then, the range of skills and knowledge required for information literacy has expanded to include newer forms of literacy to accommodate the continually developing requirements for effective information handling and to become more suitable for complex and diversified information environments (Behrens, 1994). On the one hand, scholars or practitioners divide information literacy skills differently, including basic theoretical knowledge and basic application skills of information and information technology; the ability to use information technology to study, communicate, cooperate and solve problems; and information awareness and ethics (ACRL, 2000; UNESCO, 2010). On the other hand, “computer literacy”, “media literacy”, “network literacy”, “digital literacy”, “ICT”, “data literacy”, “mobile literacy”, and other concepts have emerged as extensions of the concept of information literacy and have been widely discussed (Bruce, 2000; Pinto et al., 2010, Pinto et al., 2019; Stopar & Bartol, 2019). Although the connotation and extension of information literacy are constantly being adjusted and revised, it is always defined as a series of definable and standardized capabilities that guide people to obtain, screen, evaluate and integrate useful information from rich and diverse information sources to determine a direction of action.

Other scholars and practitioners believe that information literacy is no longer a simple universal skill but a practice activity that cannot be taught independently of the knowledge domains, organizations, and practical tasks in which these skills are used (Limberg et al., 2012; Tuominen et al., 2005). In fact, information literacy in the workplace has been a concern for a long time and is considered “a model of an information seeking and using process in the workplace” (Wai-yi, 1998). Previously, the information literacy of different roles was explored (Lloyd, 2005), but today such a research direction is relatively scarce. Fortunately, with the explosive growth of information in the field of education and health, which has posed a substantial challenge to professionals and citizens, information literacy has rapidly developed in specific work scenarios. In education, the information literacy of students (Zhu et al., 2019), teachers (Zhou et al., 2020) and education managers (Celep & Tülübaşı, 2014) has been widely examined for the promotion of education reform and the realization of education innovation. In the domain of health, many health information resources published on the Internet increasingly require people to effectively locate, critically evaluate, and efficiently use them (Haruna & Hu, 2018).

Therefore, information literacy is a concept with a long developmental time, rich connotation and extension, and wide application scenarios, making it of great significance to adapt to the vigorous development of technology. Studies on information literacy topics and their changes can be informative for researchers and can help them to better understand the domains of such topics, that is, by clarifying the developmental context of information literacy, finding weak points and grasping future trends, which will contribute to the development and innovation of information literacy. Some studies have classified the topics of information literacy via qualitative systematic reviews (Sproles et al., 2013), content analyses (Wu, Li, et al., 2020), bibliometric analyses (Kolle, 2017; Park & Kim, 2011), citation analyses (Taşkın et al., 2013) and word co-occurrence analyses (Pinto, 2015; Pinto et al., 2014, b, 2019). These studies are often focused on limited scenarios, such as higher education and information literacy education. However, the analyses of hot spots and topics are often broad, and a dynamic trend analysis has yet to be conducted.

In this paper, we conducted a topic model dynamic analysis of the articles on information literacy studies in the Web of Science (WOS) core collection database from 2005 to 2019. The global topics and their popularities, topic similarities and correlations, along with the temporal evolution of local topics and the subject diffusion of local topics were

analyzed and presented. In addition, future research directions were discussed. This study focused on three major research questions:

1. What are the research topics within information literacy?
2. How did the research topics evolve along the temporal dimension?
3. How did the research topics diffuse among the fields of subjects?

Literature review

Studies on information literacy and its changes can be informative for researchers and can enable them to better understand its domains. Researchers and practitioners usually explore the current status and trends of information literacy research in various ways.

In the early stages, information literacy standards or projects were reviewed to grasp the key points of information literacy development. Behrens (1994) analyzed the concept of information literacy by investigating some leading definitions and descriptions of information literacy and found that “educating for information literacy”, “exploring the literacy continuum” and the “role of librarians” were developmental trends. Bruce (2000) provided an overview of contemporary information research and practice and summarized various efforts to seek new directions in educational, community and workplace contexts. Marcum (2002) proposed that information literacy be refocused away from information towards learning and beyond literacy in the direction of sociotechnical fluency. At this stage, information literacy was still in its infancy, focused on the connotation and extension of the concept.

With the development of bibliometrics, scholars began to explore the development status of information literacy through the statistical analysis of existing research (Nazim & Ahmad, 2007; Panda et al., 2013; Pinto et al., 2013). In addition to the analysis of yearly statistical trends, active authors, productive institutions, and representative journals, popular research topics were identified according to keyword frequency. Pinto et al. (2010) showed that “research libraries”, “information literacy”, “information seeking”, “academic librarianship” and “literacy skills” were the most significant terms in the selected texts, illustrating how information literacy has been progressively incorporated into the library and academic fields. Aharony (2010) found three main topic types—“miscellaneous”, “health and medicine”, and “education”—which reflected a tendency to associate information literacy with health and medicine and stressed people’s need for information literacy in this specific context. Park and Kim (2011) found five major clusters calculated by the nearest neighbor cluster program, while “user training” and “students” were major descriptors in the subtopic area of information literacy, confirming that the research area has focused mostly on higher education, school libraries, and the education sector in general. Thus, information literacy research in the fields of library and information science (LIS) and education still plays a leading role and is expanding the research territory for diverse populations in communities, workplaces, and other contexts.

Next, researchers pay more attention to the topics of information literacy in specific fields or domains with effective methods, such as multidimensional scaling and network analyses. Hsieh et al. (2013) found that “information literacy”, “media literacy”, and “digital literacy” were treated significantly different between the United States and Taiwan by exploring the characteristics of theses and dissertations on information literacy from 1988 to 2010. Sproles et al. (2013) analyzed the literature on information literacy and library

instruction from 2001 to 2010 and discovered that key topics continue to be collaboration, assessment, and the application of technology to instruction efforts. Pinto et al. (2014) compared the conceptual structure of information literacy between the social sciences and health sciences from 1974 to 2011. The area of health sciences yielded four clusters, and the most central descriptor was “education”, which was strongly linked to “information retrieval” and weakly linked to “information skills”, “information seeking”, and “information science”. The social sciences had six clusters in which “information literacy” and “education” had the most occurrences. Kumar (2014) examined the scientific productivity of digital literacy in Online Library Information Science and Technology Abstracts (LISTA) from 1997 to 2011 and found that majority of articles focused mainly on academic education. Pinto et al. (2014) diagnosed the scientific production of Ibero-American researchers on information literacy from 1985 to 2013. In addition to the most common term “information literacy”, “training in information competences” has gradually gained a presence in education, informatics, and communication areas. Li et al. (2015) found that “information literacy” and “information” were the most frequent keywords in the domain of health information-seeking behavior. Pinto (2015) examined the subject area of Information Literacy Assessment in Higher Education (ILAHE) in a retrospective and selective search from 2000 to 2011. Five clusters—“evaluation education”, “assessment”, “students efficacy”, “learning research”, and “library”—were identified but overlapped significantly as a result of their terminological “proximity”. Tallolli and Mulla (2016) indicated that library and information science researchers have made significant contributions to information literacy studies. Bhardwaj (2017) showed that research on information literacy in developing countries was unpopular in the humanities and social sciences. Related studies focused on the primary disciplines showed that “information literacy”, “education” and “evaluation” remained the core terms or topics.

Influenced by social development, technological updates and field integration, related studies have also paid more attention to the future direction of information literacy topics. According to Kolle (2017), “digital divide”, “media literacy”, “pedagogy”, “higher education” and “critical thinking” were popular research topics in the IL domain from 2005 to 2014. In addition, he pointed out that the assessment of IL among students in higher education appears to be a new area of IL research. Martzoukou and Sayyad Abdi (2017) categorized the key research directions into four broad contextual areas—“encompassing leisure and community activities”, “citizenship and the fulfillment of social roles”, “public health” and “critical life situations”—which pointed to the need to develop an information literacy mindset. Stopar and Bartol (2019) showed that the computer-, information-, and digital-related terms that define these clusters predominate in different periods: the computer-related terms are earlier terms, which are followed by the information- and digital-related terms, as mature terminology is used to embrace more trendy novel concepts. They suggested that all fields should collaborate in the future. Pinto et al. (2019) offered a bibliometric analysis of the scientific production on Mobile Information Literacy in Higher Education published between 2006 and 2017. They note that incorporating IL in the context of online learning highlights how mobile technologies are gaining ground at all levels. Pinto et al. (2020) further investigated the development of mobile information literacy from 2006 to 2019, and six clusters were identified—“IL and e-learning”, “mobile devices and competencies”, “ethics”, “library and e-resources”, “educational technology” and “technological environment”—demonstrating the growing interdisciplinarity of scientific publications on “mobile information literacy”. Wu et al. (Wu, Li, et al., 2020) conducted source tracing, an association analysis and feature mining of the research content of information literacy theory from five aspects—“placement”, “dialogue”, “application”, “evaluation” and

“generation”—emphasizing that information literacy should be interdisciplinary, should deepen theory and should be people-oriented.

Overall, the research topics of information literacy diversified over time, and interdisciplinary integration was encouraged. However, those studies remain targeted towards limited scenarios, such as higher education and information literacy education. In addition, the analyses of hot spots and topics are often broad, and a dynamic trend analysis has yet to be conducted. Therefore, it is necessary to further retrospectively examine the topical development of information literacy from a dynamic perspective. On the other hand, those studies have classified the topics of information literacy via qualitative systematic review content, bibliometric, citation and word co-occurrence analyses. However, with the development of technology and the differences in domain characteristics and user requirements, the corresponding information literacy studies have a limited focus, which is often ignored in the analysis of high-frequency words. It is highly necessary to combine the textual content (such as titles and abstracts) to obtain a deeper classification, which will facilitate understanding of the topical development of information literacy studies. For topic detection and tracking, topic modeling is regarded as more flexible and effective than alternative approaches such as document clustering (Kuhn, 2018). A word-embedding algorithm can be used to represent words semantically based on the context to calculate topic correlations (Chen et al., 2017; Xie et al., 2020). Therefore, this study attempts to use LDA, which is a widely used topic modeling tool, to describe the topics of information literacy and to explore the evolution and diffusion of these topics in combination with the BERT word-embedding algorithm.

Data and methodology

Figure 1 presents a research overview, which illustrates the dataset acquisition and analysis methodology. The overall scheme can be divided into three subprocesses: (1) data retrieval and preprocessing, (2) topic modeling, and (3) topical evolution and diffusion. Meanwhile, the analytical process of this paper is based on the Python environment, and some

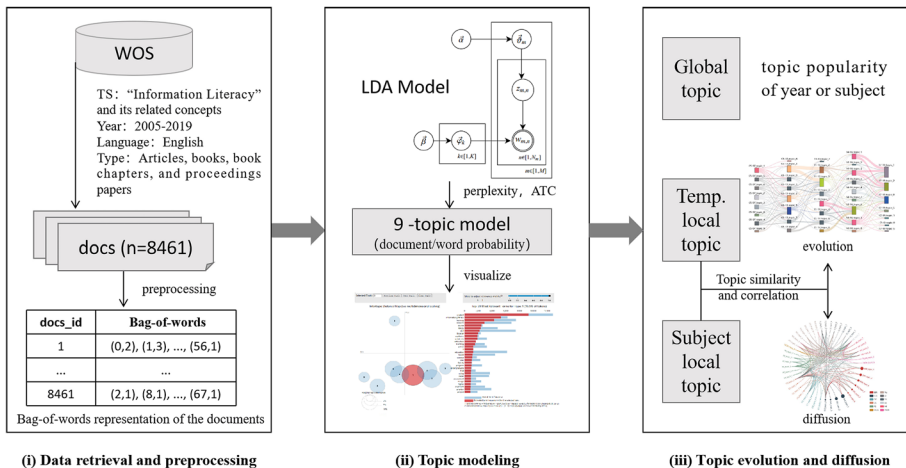


Fig. 1 Research overview

open-source textual analysis algorithms and data visualization tools are used to analyze or visualize the results.

Data retrieval and preprocessing

The nature of our study required us to identify the search terms, as there are related or similar concepts. On the one hand, the concept is not limited to a single category of information; related terms that are similar to “computer”, “network”, “media”, “ICT”, “digital”, “data” and “mobile” have emerged based on the background of technological development. On the other hand, literacy, competence and skill are often used in the context of these terms. Scientific publication data that were related to information literacy and its related concepts were gathered via a search strategy from the WOS core collection database in June 2020. Based on the above, a Boolean search equation was applied to expand the search to related concepts:

TS= (“computer competen*” OR “computer literac*” OR “computer skill*” OR “data competen*” OR “data literac*” OR “data skill*” OR “digital competen*” OR “digital literac*” OR “digital skill*” OR “ICT competen*” OR “ICT literac*” OR “ICT skill*” OR “information competen*” OR “information literac*” OR “information skill*” OR “media literac*” OR “internet literac*” OR “meta literac*” OR “mobile literac*” OR “Information Communications Technology competen*” OR “Information Communications Technology skill*”)

The selected document types included articles, books, book chapters, and proceedings papers. Based on the search results, only a few related papers were published before 2005, so we finally collected 8461 English-language documents published from 2005 to 2019. The distributions of the publication years and the main subjects are shown in Fig. 2. These distributions show the developmental trend and multidisciplinary of this field. In 2015, the number of documents increased significantly, and the documents involved 128 subjects (WOS research areas), which included 12 main subjects that were covered by more than 200 published papers. The percentage of papers that covered these subjects was 97.70%. The researchers are mainly in the fields of education and educational research (EER) and

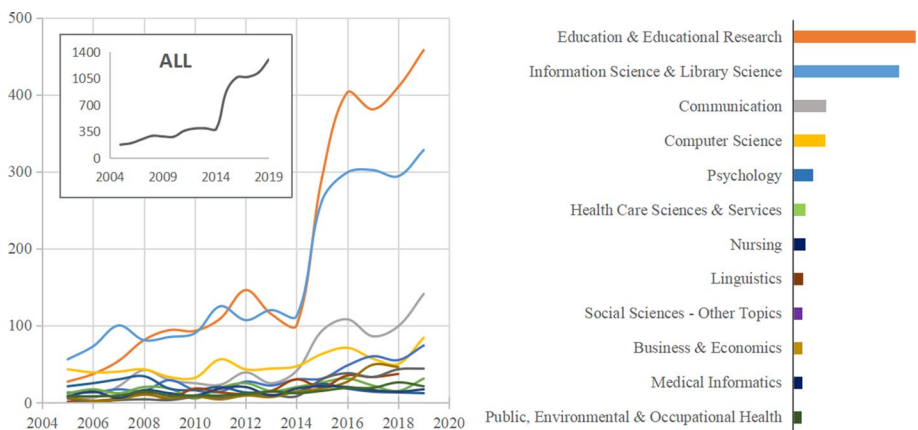


Fig. 2 Distributions of the publication years and the main subjects

information science and library science (ISIS) for the following reasons: (1) information literacy was developed from and has become an important integral part of ISIS, and, (2) with the rapid development of educational technology, the information literacy and other related digital competences of teachers and students have received substantial attention.

To increase the data quality, preprocessing was conducted by using the following Python package: Natural Language Toolkit (NLTK) (Bird et al., 2009). First, the title and the abstract were used as the text corpus to extract the terms. Second, numbers, punctuation symbols, and stop words were deleted. Third, all terms were converted to the singular form and the lower case and were stemmed to produce more readable words. Fourth, bigrams that appeared 20 times or more were added to documents. Fifth, we filtered out words that occurred in less than 20 documents or in more than 50% of the documents and saved these as a dictionary. Finally, we transformed the documents to a vectorized form by simply computing the frequencies of the words in the dictionary.

Topic modeling

Latent Dirichlet allocation (LDA), which was proposed by Blei et al. (2003), was used to extract topics from the corpus. LDA is a probabilistic model in which each document in a corpus is described by a random mixture over latent topics. Each of the latent topics is characterized by a distribution over words. The Gensim library (Rehurek & Sojka, 2010) was utilized for LDA modeling.

As LDA parameters, the perplexity and average topic coherence (ATC) were estimated within Gensim for the selection of the number of topics (as presented in Fig. 3). The perplexity is only a crude measure; it is helpful for getting close to a suitable number of topics in a corpus. The average topic coherence, which measures the degree of semantic similarity between high-scoring words in the topics, can help distinguish between topics that are semantically interpretable and topics that are artifacts of statistical inference (Stevens et al., 2012). After comprehensive consideration, a 9-topic model was finally selected.

LDA modeling yielded two matrices: a matrix of document probabilities, which contains the probability distribution of each document that belongs to one of the topics, as presented in Table 1, and a matrix of word probabilities in the corpus and their association with each topic, as presented in Table 2.

In addition, the pyLDAvis package (Sievert & Shirley, 2014) can be used to visualize topics. An intertopic distance map, which was obtained via multidimensional scaling, is presented on the left, and the top 30 most relevant words and their frequencies for each topic are presented on the other side.

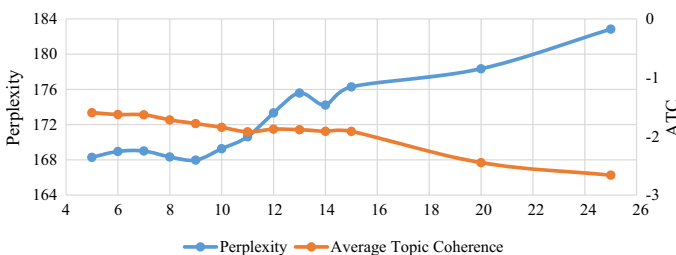


Fig. 3 Perplexity and average topic coherence for an n-topic model

Table 1 Document probabilities

docs_id	year	subject	topic 1	topic 2	topic 3	topic 4	topic 5	topic 6	topic 7	topic 8	topic 9
1	2005	ISIS	0.2562	0.2083	0.1252	0.0457	0.1700	0.0000	0.0000	0.2562	0.0000
2	2005	EER	0.4194	0.1798	0.0000	0.0648	0.2233	0.0000	0.0000	0.4194	0.0000
...
8641	2019	EER	0.0000	0.0851	0.0493	0.1372	0.0000	0.0000	0.0000	0.0000	0.0000

Table 2 Word probabilities

words	topic 1	topic 2	topic 3	topic 4	topic 5	topic 6	topic 7	topic 8	topic 9
Student	0.0476	0.0000	0.0002	0.0215	0.0000	0.0000	0.0000	0.0000	0.0112
Library	0.0120	0.0159	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
...

Topical evolution and diffusion

We explored the developmental paths of information literacy topics from the temporal and subject perspectives. To investigate the temporal dynamics, the corpus was divided into five three-year time spans: 2005–2007, 2008–2010, 2011–2013, 2014–2016, and 2017–2019. The topic models were trained individually for each time span. To identify the topical relationships between the main subjects, each subject topic model was estimated. For differentiation, we refer to the 9 general topics as global topics, the time-span topics as temporal local topics, and the subject topics as subject local topics.

Global topic popularity

Global topic popularity was calculated by aggregating the per-document topic distribution by year or subject dimension. First, the probabilities of the same topic in the same dimension were summed to obtain the total probability. Next, the results were normalized by the number of publications in the dimension. Finally, we obtained the global topic popularity values for additional years or subjects using visualization tools. When the proportion is higher, the topic is more popular. In addition, a graphing tool named Cluster Purity Visualizer (Swamy, 2016) was implemented to obtain a basic distribution graph in different subjects.

Topical similarity

A topical similarity formula established by (Xie et al., 2020) was extended for calculating the multiperiod and multisubject topical similarity as follows:

$$\text{Sim}(T_i^N, T_j^M) = \sum_{k=1}^K P_k^n(T_{i,k}^N) \times P_k^n(T_{j,k}^M) \times S_{k(N,M)}(T_{i,k}^N, T_{j,k}^M)$$

where T_i^N is topic i in topic N , $T_{i,k}^N$ is the topic word in topic i , T_j^M is topic j in topic M , $T_{j,k}^M$ is the topic word in topic j , K is the number of topic words in a topic, and p_k^n is the normalized value that is obtained via formula $1 - \log(p_k)$, in which p_k is the LDA probability value. Then, $S_{k(N,M)}(T_{i,k}^N, T_{j,k}^M)$ is the averaged tensor similarity between $(T_{i,k}^N, T_{j,k}^M)$, which is obtained by the embedding algorithm.

We selected BERT (Devlin et al., 2019) as the representation algorithm. The BERT algorithm is a new language representation model that stands for bidirectional encoder representations from transformers. In contrast to word2vec, BERT is designed to pretrain deep bidirectional representations from unlabeled text by jointly conditioning both the left and right contexts in all layers. The BERT base uncased with 12 layers, 768 hidden nodes, 12 heads, and 110 M parameters from the Google BERT site was used for superior training. First, sentences were separated by separator [CLS] at the front and separator [SEP] at the end of each sentence. Then, three embeddings were calculated for each sentence: token, segment and position. The input of BERT was formed by superimposing the three embeddings. Finally, we obtained tensors for various sentences. The tensor of a word may differ among contexts. Therefore, the averaged tensor was calculated based on the word-corresponding sentences in the local range for a target topic word. Figure 4 presents an example of the calculation process for word $T_{1,1}^G$, which denotes the first word, namely, “student”, in global topic 1. Then, the final tensor is compared with the word “student” in global topic 9.

Documents that contained the word “student” numbered 3598, 994 of which belonged to global topic 1 and 22 of which belonged to global topic 9. In topic 1, 3670 sentences and their tensors were further extracted by using the BERT pretrained model. Each sentence was represented by a vector with 768 dimensions. Then, the average tensor was used as the final tensor for the word “student” in topic 1. The “student” final average tensor in topic 9 was calculated via the same method. The high-dimensional tensors of the “student” average values (in the red rectangle) and 20 randomly selected sentences in each topic (pink for topic 1 and blue for topic 9) were visualized by the Embedding Projector. The meaning of “student” differed among topics, and the semantic relationships among sentences were closer under the same topic. Therefore, this method can effectively identify the direct similarities of topic words.

Topical correlations

Based on all final topic similarities in the corresponding context, we examined the topic similarity average value and the products of this value and 0.8, 0.9, 1.1, 1.2, 1.5, 2, etc. To identify the correlations between topics as clearly as possible, we set two thresholds. If the similarity score between two topics exceeded the corresponding threshold, we defined the relation as topic correlation. The average value was set as the threshold for temporal local topics, while 1.2 times the average value was used for subject local topics.

Evolution of temporal local topics

We named each temporal local topic based on the temporal span and proportion ranks. For example, the topic with the largest proportions from 2005 to 2008 was named 05-11_topic_1. We calculated the similarities between the temporal local topics and their subsequent temporal local topics, and we identified the correlation pairs until all periods had been considered. Then, the evolutionary paths of all topics were multiplied to obtain the

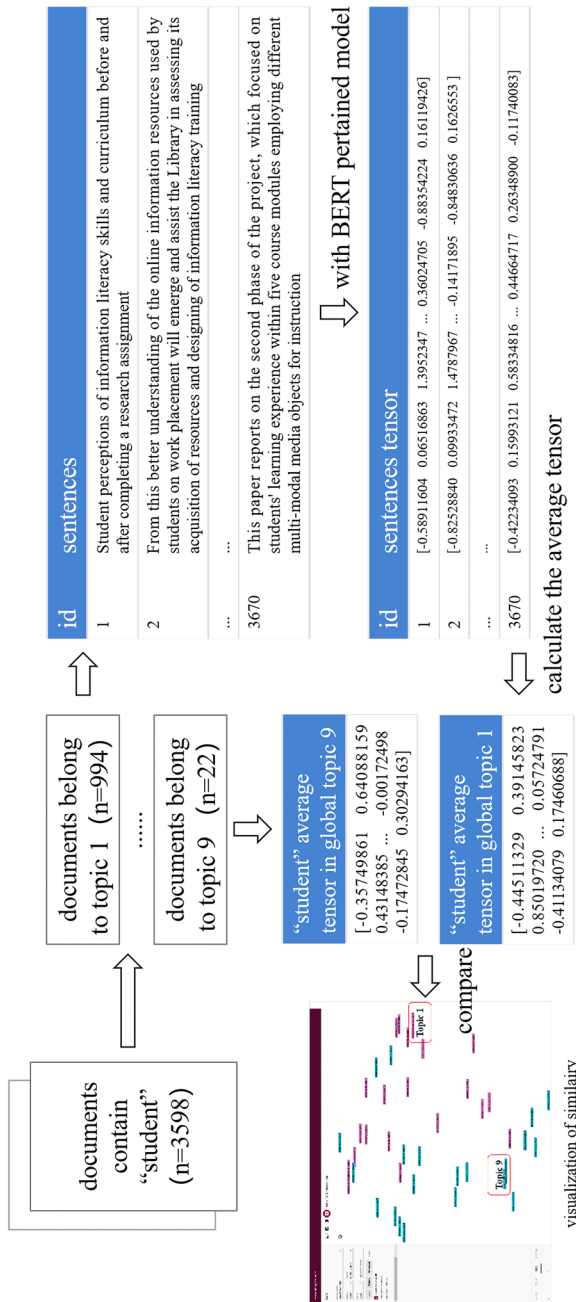


Fig. 4 Calculation process for the word "student". (Color figure online)

evolutionary path of information literacy topics by using the Sankey diagram as a visualization tool with Apache ECharts (Li et al., 2018).

Diffusion of subject local topics

We named each temporal local topic based on the subject and proportion ranks. For example, the topic with the largest proportions in education and educational research (EER) was named EER_topic_1. The similarities between topics in different subjects were calculated to obtain all topic correlations. Next, all topic pairs for the increasing year (cumulative to 25%) of subject local topics were multiplied to obtain the diffusion of information literacy topics by using chord diagrams in Apache ECharts.

Results

Global topics

The results for the global topics are presented in Table 3. For each topic, Table 3 lists the topic_id, the suggested topic labels, the top 5 words with the highest probabilities, and the topic proportions in the complete corpus.

Nine topics resulted—“learning and education”, “library service”, “new digital technology”, “teacher ICT”, “health information”, “Internet use”, “medium literacy”, “evaluation”, and “computer skill”—which were suggested by two domain experts. The relationship among the topics is illustrated in Fig. 5.

Topic 1, which has the largest share and the most related topics, focuses on student information literacy learning. It is the original objective of information literacy development and an important issue in the information or intelligence age. Topic 2, which concerns library service for information literacy, ranked second. Libraries have long adopted the important responsibility of information literacy education due to their rich information resources, advanced information technology, professional training teams, and suitable learning environments. Topic 3, which also occupies a relatively large share, considers information literacy practice from a broader perspective and focuses on the objective of keeping pace with the current technology, such as digital technology. Topics 4 and 5 consider information literacy for two groups: teachers and patients, respectively. Education and health are both important application domains of information literacy. Topics 6–9 occupy smaller shares and are independent of other topics. Topic 7 is related to medium literacy, of which the scope is information in various media, while the objective is the same as that of information literacy: digital survival. The other topics are related to information literacy measurement: topics 6 and 9 are the dimensions, while topic 8 corresponds to evaluation research.

Global topical popularity

The global topical popularity of years is presented in Fig. 6. “Learning and education” and “library service” were relatively stable and have many papers, which together constitute a share of nearly 40%. “New digital technology” and “teacher ICT” showed increasing trends, each with 10% growth, while the shares of “health information” and

Table 3 Results for the global topics

Topic_id	Label	Words of the topic	Proportion (%)
1	Learning and Education	0.048*student + 0.027*information_literacy + 0.022*learning + 0.015*research + 0.013*course	19.5
2	Library Service	0.016*library + 0.013*data + 0.011*paper + 0.011*research + 0.010*service	15.5
3	New Digital Technology	0.044*digital + 0.014*social + 0.011*practice + 0.010*learning + 0.009*technology	15.0
4	Teacher ICT	0.039*teacher + 0.028*ict + 0.023*school + 0.022*student + 0.019*competence	13.7
5	Health Information	0.038*health + 0.015*use + 0.015*patient + 0.011*web + 0.010*care	11.5
6	Internet Use	0.019*internet + 0.013*use + 0.012*group + 0.011*intervention + 0.010*adult	9.1
7	Medium Literacy	0.098*medium + 0.028*medium_literacy + 0.026*education + 0.023*critical + 0.014*social	5.8
8	Evaluation	0.017*analysis + 0.015*test + 0.015*data + 0.013*scale + 0.011*assessment	4.9
9	Computer Skill	0.083*computer + 0.043*skill + 0.024*self + 0.017*nursing + 0.015*internet	4.8

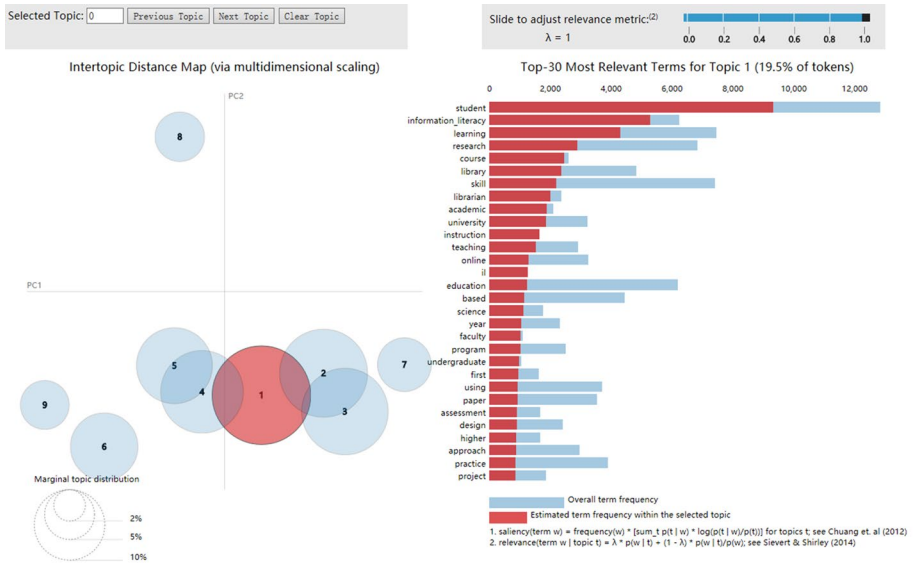


Fig. 5 Global topics with pyLDAvis

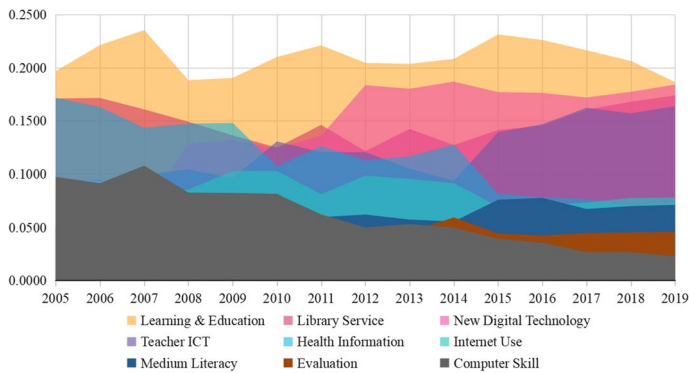


Fig. 6 Global topical popularity of years

“computer skill” halved. The remaining topics—namely, “Internet use”, “medium literacy”, and “evaluation”—underwent generally mild trends and had relatively low shares.

The global topic popularities of the main subjects are presented in Fig. 7. The topic preference differs among the subjects. More topics are covered in the fields of education and educational research (EER); communication (CO); computer science (CS); health care science and services (HCSS); nursing (NU) and public, environmental and occupational health (PEOH). In EER, multiple topics show increasing trends. In CS, the popularities of the top five topics are relatively average; hence, the research content of information literacy in the CS field is wider. In each of the remaining subjects, studies focused on a single topic; for example, in information science and library science (ISIS), studies concentrated on “learning and education”, and, in psychology (PS), studies focused on “Internet use”.

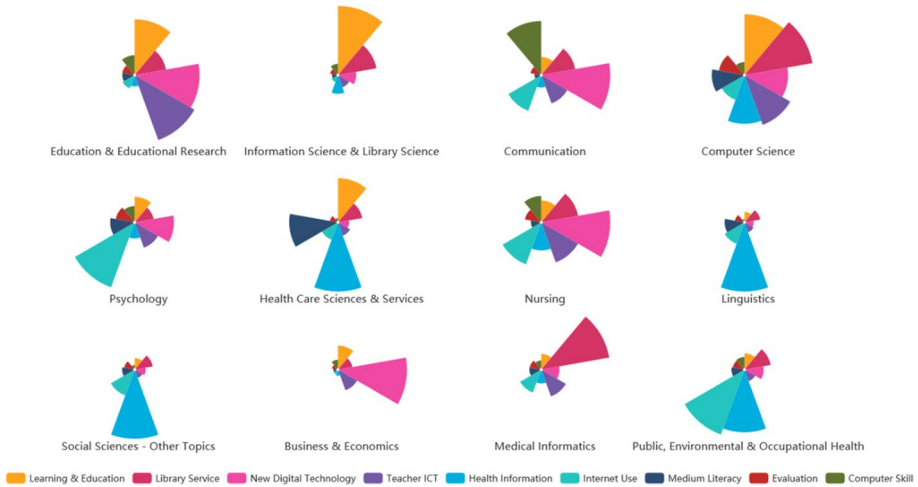


Fig. 7 Global topical popularities of the main subjects

Temporal local topics

The top five words of all temporal local subjects were statistically classified into six categories: ability, technology, field, people, place and application. The main words and frequencies of each part are listed in Table 4.

First, most of the key words still focus on the embodiment of ability and technology for the field, people and place. In addition to the emergence of media and the Internet, various changes have occurred in technology. Earlier words were mainly “computer”, “online” and “web”, which were related to the main technologies that appeared at that time. In recent years, “digital”, “social” and “data” have emerged, which have posed new challenges to people’s information literacy. Moreover, field, people, and place were mainly concentrated in education and health and were affected by the informationized development of these two fields. Unfortunately, there was no special bias in the application, which had not yet formed a relatively mature and rich application of the research dimension.

Evolution of information literacy topics

According to the similarities between the temporal local topics, the topical evolution from 2005 to 2019 is presented in Fig. 8. An emerging topic refers to a topic that had weaker or no correlations in the previous period, such as 11-13_topic_8. A disappearing topic is one that has weaker or no correlations in the subsequent period, such as 08-10_topic_6 and 11-13_topic_7. In addition, some topics appear only in a single period of time and subsequently disappear, such as 05-07_topic_7, 08-10_topic_8, 11-13_topic_5, 17-19_topic_4, and 17-19_topic_7; they are not shown on the evolutionary path. The topics that are discussed above mainly refer to two aspects, thereby reflecting two evolutionary mechanisms:

Table 4 Words in temporal local topics

Type	Topic words	2005–2007	2008–2010	2011–2013	2014–2016	2017–2019
Ability	skill (12), information_literacy (9) ict (5)					
Technology	literacy (5), medium_literacy medium (9), internet (7) computer (8)	digital_literacy		–	medium_literacy, competence	medium_literacy
Field	technology –	digital (8) online, web		technology (5) social	social, data, online, social_ medium	social (2), data
People	health (9), learning (7) education, nursing student (19)	education, nursing (2)		–	education (2), news	education (2), news (2)
Place	librarian – library (8) university	child (5), teacher (4) librarian patient, adult		patient, adolescent	librarian –	librarian patient, girl
Application	use (16), research (8) practice, program, self, approach, anxiety, system, training	school (2) practice, program, paper, search		school practice, program, develop- ment, care, access, based, article, intervention	school, academic support, critical, knowledge, test, based, reading, article	– practice, method reading, pre, image, result, self, body, group, test, development

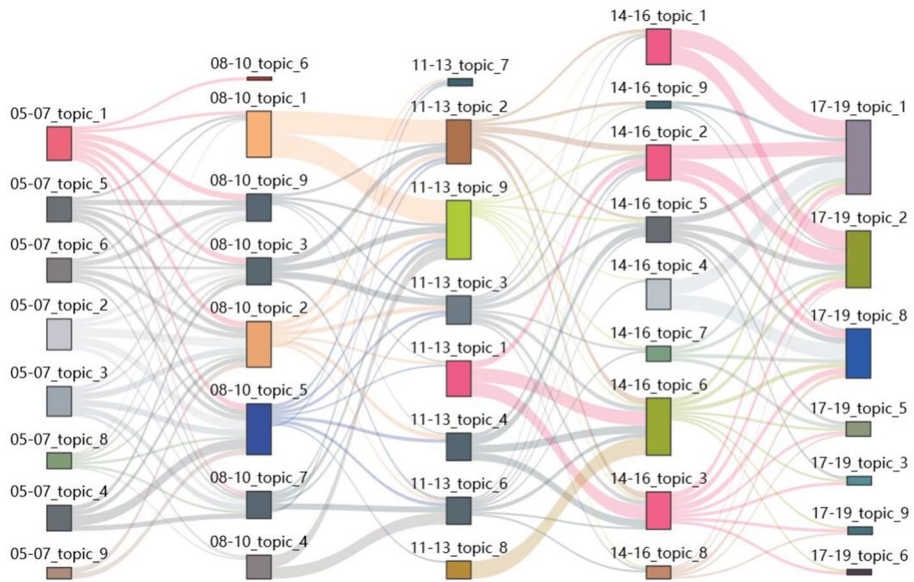


Fig. 8 Evolution of temporal local topics

Transferring

Several topics contain “medium”; their similarities were slightly lower due to their differences in context. In 05-07_topic_7 (medium, literacy, medium_literacy, use, and system), medium literacy was emphasized. In 08-10_topic_8 (medium, adult, child, digital, and literacy), children and adults were emphasized. In 11-13_topic_8 (medium, child, intervention, program, and adolescent), young people’s abilities were developed through intervention or projects. In addition, emerging topic 17-19_topic_7 (child, test, group, use, and reading), in which there was no mention of “medium”, focused on children’s media literacy in the smaller scope of “reading”. The focuses of these topics vary; hence, the evolutionary path shows no continuity.

Suspending

Here, topics on “health” were considered. The skills of nursing students were considered in 8-10_topic_6 (student, program, skill, health, and nursing). Both 11-13_topic_7 (patient, health, care, access, and based) and 17-19_topic_4 (health, patient, method, skill, and student) attached importance to evaluation, but consideration of these topics was suspended from 2014 to 2016. These topics have received renewed attention.

Most of the topics were in the process of continuous merging and splitting, thereby leading to the dynamic evolution of the research topic focus. Two evolutionary mechanisms were identified:

Crossing

Crossing typically occurs on topics with widely used words. These topics had a wide range of meanings, with many sources and divisions, such as 05-07_topic_1 (student, learning, literacy, information_literacy, and education), 08-10_topic_2 (medium, education, learning, practice, and student), 11-13_topic_9 (computer, use, student, technology, and learning), 14-16_topic_6 (teacher, ict, competence, education, and school), and 17-19_topic_1 (student, teacher, learning, education, and digital). These topic words were mostly core high-frequency words, which easily correlate with other topics. However, these topics and evolutionary pathways are too broad to reflect knowledge transfer.

Growing

This term refers to the relatively limited and stable mechanism of topical change in evolution. The number of related topics should be controlled within a limited range; an example is presented in Table 5, which corresponds to the dynamics of 08_10_topic_4 and 14_16_topic_4. 08_10_topic_4 was the combination of health and computer-related skills to form the content of health online search, which was divided into health and computer in subsequent studies. Compared to the words in 2005–2007, the health and computer research content changed in 2011–2013: 11-13_topic_6 placed more emphasis on technology (such as “internet” and “computer”), and 11-13_topic_9 emphasized “learning”. Then, these merged into 14-16_topic_4, which considered the use of health data. This topic was related to data and was influenced by the development of big data. However, the related topics in 2017–2019 did not contain the word “health” but had wide ranges. This finding may be due to the smaller amount of research on data analysis in the health domain. Hence, 14-16_topic_4 can be divided only into broad topics. However, in any case, these topics continue to grow as the environment changes.

In summary, crossing was the main evolutionary mechanism in the temporal local topics of information literacy. The core research topics—such as learning, library and technology—were all considered, but the focus changed according to the background of the times. However, the core topic words were relatively stable, and unfortunately, few new research directions in information literacy have been investigated in recent years.

Table 5 Dynamics of 08-10_topic_4 and 14-16_topic_4

2005–2007	2008–2010	2011–2013	2014–2016	2017–2019
topic_3: skill_use_ health_nurs- ing_training	topic_4: health_ online_internet_ search_skill	topic_6: health_ internet_skill_ use_computer	topic_4: use_data_ health_technol- ogy_internet	topic_1: student_ teacher_learn- ing_education_ digital
topic_8: computer_ self_use_stu- dent_internet		topic_9: com- puter_use_stu- dent_technol- ogy_learning		topic_8: medium_ medium_liter- acy_news_child_ education

Subject local topics

The top five words of all subject local topics were statistically classified as described above. The main words and frequencies of each part are listed in Table 6. On the left are words that appear in multiple subject local topics, and those words cover more than one-third of the subjects. On the right are words with fewer subjects; there were 16 words in three subjects, 36 words in two subjects, and 74 words in only one subject.

High-frequency words have also appeared in various subject local topics. However, many words still only appeared in a few subjects. For example, in terms of technology, social media was more discussed in psychology (PS) and social science (SS) regarding the use of various types of users on social media. Health information was emphasized in health care science and service (HCSS) and public, environmental and occupational health (PEOH), which refers to a special type of information. Web was also used in two subjects on influencing factors around related intervention experiments. Application had more words that appeared only a few times because each field has its own characteristic application aspects. For example, care was used in three health-related subjects (HCSS, NU, and MI), and self was related to measurements, which were more common in EER, ISIS and PS.

Diffusion of information literacy topics

The topical diffusion of subject local topics among 12 main subjects is illustrated in Fig. 9. Among the 108 subject local topics, 47 (43.52%) were related to other topics. Most of the EER, PS, and NU topics were related to other subject topics, while only one or two local topics in HCSS, BE and PEOH appeared in the diffusion diagram.

Table 7 lists and describes the nine topics that have extensive links to other topics. CS_topic_4 (student, ict, teacher, school, and learning), which has 33 links with other topics, has two core words that correspond to ICT: teachers and students. Similarly, the contents of other topics were core words in various subjects, which mainly included students' learning, the use of a computer or network, and children's skills, among others. Most of these topics were developed prior to 2010. The earliest was NU_topic_2 (technology, learning, student, education, and communication). This topic corresponds to student education in the field of nursing. The nursing field has focused on information literacy education for a long time due to the development of health informatization.

In addition, this paper examined the diffusion of topics between each pair of subjects and identified the diffusion mechanisms, as presented in Table 8.

Absorbing

This term refers to the scenario in which there was only a single topic connection between two subjects and there was an order of the rising year. This action supports an absorption relationship of the topic between subjects. Based on absorbing, *absorbing with division* and *absorbing with merger* expanded the number of topics. The former refers to many topics of one subject having absorbed another subject, while the latter refers to topic absorbance by another discipline. Here, we consider the subject communication as an example, as illustrated in Fig. 10.

Table 6 Words in subject local topics

Type	Topic words	4 > Number of subjects > 1
	Number of subjects > = 4 (1/3 of all)	
Ability	digital (9), skill (9), ict (6), medium_literacy (5), competency (5)	literacy (NU, BE, MI), competence (EER, LI, SS), information_literacy (EER, CS)
Technology	technology (10), internet (10), computer (8), social (7), data (7), medium (6), system (6), online (6), information (5)	social_medium (PS, SS), health_information (HSCC, PEOH), web (ISIS, NU)
Field	learning (7), education (7), health (5)	nursing (CS, BE, PEOH), reading (ISIS, PS, LI), clinical (HCSS, NU, MI), eHealth (HUCC, MI), news (CO, SS)
People	student (11), child (9), user (7), patient (6), teacher (6)	adolescent (CO, MI, PEOH), participant (HCSS, MI), parent (CO, PS), young (CS, SS), nurse (HCSS, MI)
Place	school (4)	library (ISIS, CS, CO), university (ISIS, BE)
Application	research (7), practice (6), use (5), intervention (5), based (5), access (4)	care (HCSS, NU, MI), text (EER, LI), program (PS, HCSS, PEOH), self (EER, ISIS, PS), search (ISIS, CS, MI), evidence (NU, MI), paper (ISIS, CS, LI), knowledge (ISIS, PS, BE), efficacy (EER, ISIS, PS), analysis (ISIS, CO, SS), course (EER, ISIS, MI), effect (CO, BE), etc

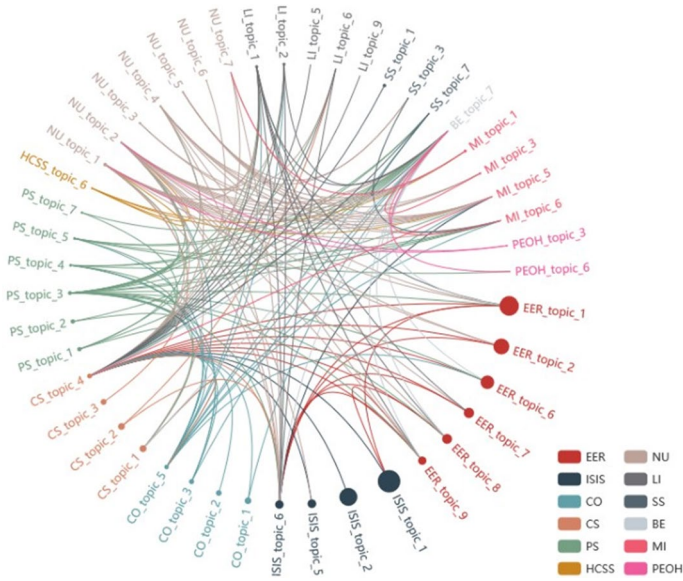


Fig. 9 Diffusion of subject local topics

Table 7 Topics with extensive links to other topics

Subject topic	Words of the topic	Number of links	Rising year
CS_topic_4	student, ict, teacher, school, learning	33	2009
BE_topic_7	literacy, advertising, training, effect, child	25	2008
NU_topic_2	technology, learning, student, education, communication	21	2006
PS_topic_3	computer, self, efficacy, self_efficacy, ict	18	2008
ISIS_topic_6	librarian, library, faculty, academic, university	16	2011
CO_topic_5	internet, use, child, skill, parent	13	2013
NU_topic_1	student, literacy, health, practice, evidence	13	2008
LI_topic_1	learning, practice, technology, teacher, student	12	2013
MI_topic_5	informatics, competency, skill, readiness, emr	12	2008

Table 8 Diffusion mechanisms between the subject local topics

Diffusion mechanism	Quantitative characteristic	Year characteristic	Number of pairs
Absorbing	One to one	Time-sequenced	11
Absorbing with division	One to all/all to all	Time-sequenced	22
Absorbing with merging	All to one	Time-sequenced	6
Paralleling	One to all	Time-paralleled	2
Interlacing	All to all	Time-interleaved	3

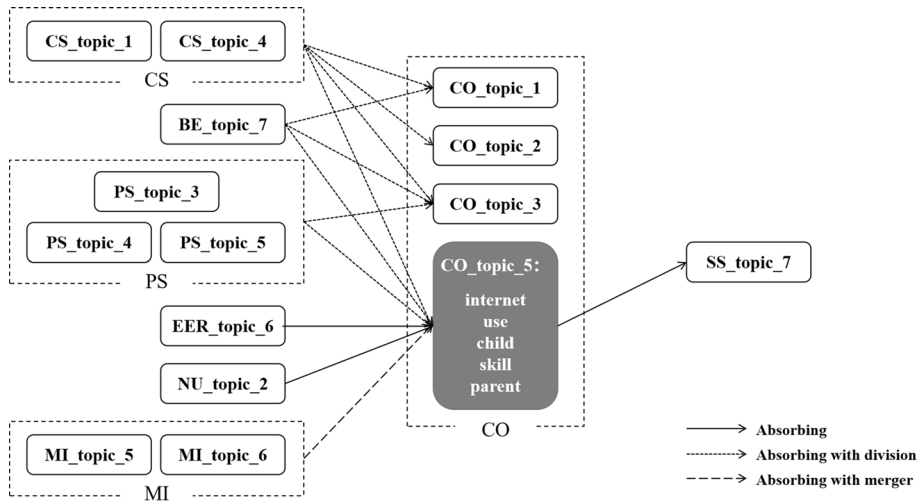


Fig. 10 Diffusion mechanisms between communication local topics and other topics

Four topics in communication (CO) have diffusion relations with other subjects. First, CO_topic_5 (internet, use, child, skill, parent) corresponds to Internet usage in the home, for which the rising year was 2013. It absorbed EER_topic_6 (computer, self, ict, level, and efficacy) and NU_topic_2 (technology, learning, student, education, and communication) and was absorbed by SS_topic_7 (medium, young, child, news, and sexual). Several CO topics absorbed topics in CS, PS and BE with the diffusion mechanism of absorbing with division, namely, CS_topic_1 (computer, digital, use, ict, and student), CS_topic_4 (student, ict, teacher, school, and learning), PS_topic_3 (computer, self, efficacy, self_efficacy, and ict), PS_topic_4 (use, child, internet, online, and parent), PS_topic_5 (medium, girl, body, medium_literacy, and intervention), and BE_topic_7 (literacy, advertising, training, effect, and child). In addition, MI_topic_5 (informatics, competency, skill, readiness, and emr) and MI_topic_6 (patient, older, adult, older_adult, and intervention) were absorbed by CO_topic_5 with merging. These topics focused on children’s or students’ computer, Internet and media usage and self-efficacy, but the characteristics of these subjects differed.

Paralleling

In paralleling, the rising years of similar topics in two subjects are the same. *Interlacing* refers to the multiple interactions between two subjects without a temporal sequence. The diffusion mechanisms of medical informatics (MI) are presented in Fig. 11 as an example.

There was a parallel relationship between MI and HCSS, while there was an interlacement with PS. Such linkages were often influenced by other common foundations and developed in a common direction. For example, MI_topic_1, MI_topic_5 and HCSS_topic_6 were associated with NU_topic_1 (student, literacy, health, practice, and evidence) and NU_topic_2 (technology, learning, student, education, and communication). These results were from an initial study of students’ health literacy and technology learning and subsequent studies of professional health technologies. Other associations between MI and PS were more complex. In addition to absorbing and being absorbed, these topics

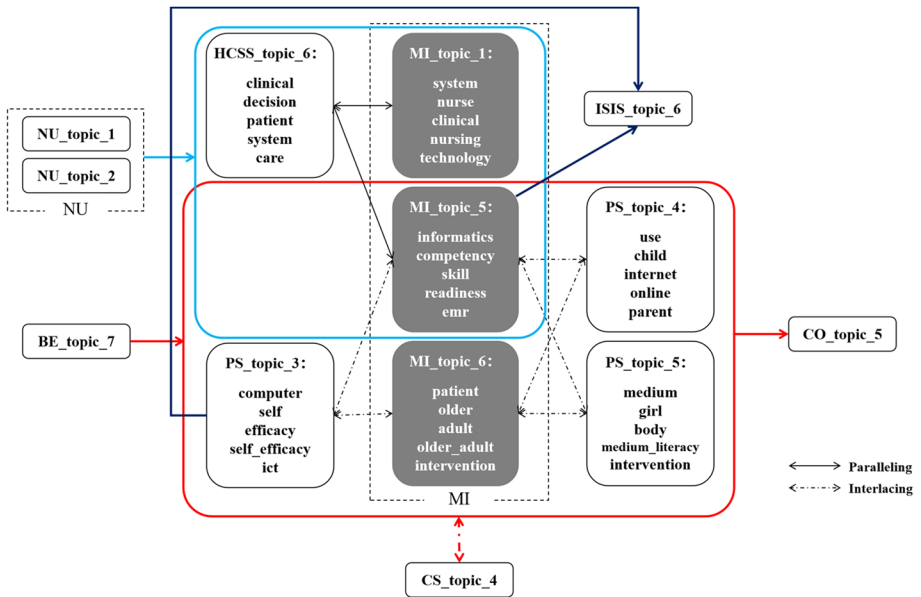


Fig. 11 Diffusion mechanisms between medical informatics local topics and other topics

also acted with a diffusion mechanism with other subjects in time. ISIS_topic_6 and CS_topic_4, which contained core words in topics, were easily connected with MI and PS, absorbed their relevant content and influenced it. Overall, MI was closely related to many health-related subjects and was developing in the direction of measurement based on psychological interaction.

First, based on the diffusion mechanisms between the subject local topics, NU had the most diffusion mechanisms with other subjects, which mainly involved “student”, “computer/internet”, “health/care” and “evidence_based” topic words. Only in the field of nursing did studies in NU begin to consider the evaluation of students’ information literacy very early. NU has formed the research scope of health-related information literacy with MI, HCSS and PEOH. In addition, it has provided ideas for the evaluation of many subjects. Second, PS, CS and ISIS presented a state of transmission, which absorbed or had been absorbed by several subjects, namely, CS for student ICT, PS for usage of technology, and ISIS for library. These subjects were often interdisciplinary; hence, the related research presented complex diffusion mechanisms. Third, EER, CO and LI absorbed the topics of related subjects; however, they were less absorbed by other subjects due to their specificity. For example, in addition to the common topic words, EER corresponded to “education”, “teacher” and “school”; CO corresponded to “social”, “news” and “message”; and LI was associated with words such as “text” and “English”. Finally, the interaction of SS and BE was relatively weak, and their research scopes were relatively small.

Overall, there were no readily observable differences in the core topics of various subjects, but preferences were identified according to the subject characteristics. There were five diffusion mechanisms of subject local topics, among which absorbing with division and absorbing were the main mechanisms, which supported the diffusion progress of information literacy studies among subjects. NU developed earlier; PS, CS and ISIS were multi-disciplinary; and EER, CO and LI showed research preferences.

Discussion

An analysis of above results reveals potential issues and directions that may require more attention and effort from scholars by focusing on the latest research on information literacy.

Focusing on the impact of information technology

The studies on information literacy presented strong “technology-related” characteristics. Three global topics were related to information technologies. The discussion on “new digital technology” was relatively more extensive and growing compared to those on “Internet use” and “computer skill”. Regarding the words in the local topics, in addition to “technology”, “computer”, “internet”, and “digital”, technology-related words often appeared, such as “medium”, “social”, “online”, and “data”. This occurrence is because information literacy is the basic literacy for citizens adapting to the information age or even the intelligent age.

Meanwhile, information technology has been developing rapidly. Figure 12 shows the search trends for the words “technology”, “computer”, “internet”, “big data”, and “artificial intelligence” in the past ten years. These trends are similar to the extension of the concept of information literacy. Bawden (2001) expanded “literacy” to include newer forms of literacy, which were more suitable for complex information environments, such as “media literacy”, “computer literacy”, “library literacy”, “network literacy”, and “internet literacy”; these types of literacy are based largely on selected skills. Furthermore, based on the background of big data, the Association of College and Research Libraries (ACRL) has used the concept of “data literacy” many times (Prado & Marzal, 2013). Pinto et al. (2019, b) adopted the concept of “mobile information literacy” for ubiquitous learning, connectivism, and multimodal learning. Thus, technology has always been an important factor in the development of information literacy.

In many studies, technology is used only as the background for exploring the influencing factors of technology acceptance (Ketikidis et al., 2012; Scherer et al., 2019), integration (Farjon et al., 2019) or use (Hughes et al., 2020) for specified people. Fewer information literacy studies focus on the impact of information technology. Technology plays an intermediary role. The deep integration of technology and other fields can contribute to a more convenient, more intelligent and fairer social environment. According to Aljanabi & AL-Hadban (2018), information literacy can strengthen learners’ satisfaction via social networking technology use. Califf and Brooks (2020) found that literacy facilitation reduces the effects of techno-stressors.

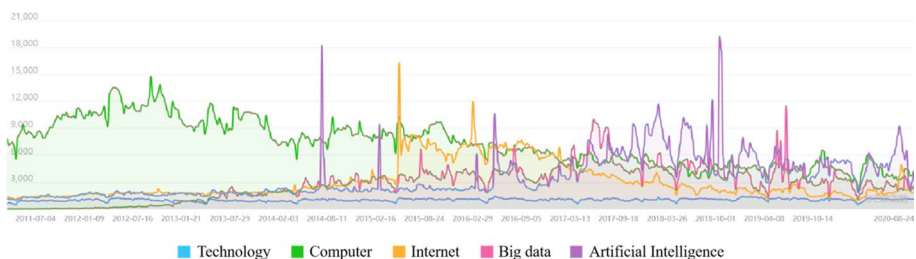


Fig. 12 Search trend chart of technical keywords

Increasing the breadth and depth of the research field

Studies on information literacy have focused on education and educational research (EER) and information science & library science (ISIS). Among the research fields, education (such as “learning” and “education”) and health (such as “health”, “nursing”, “clinical”, and “eHealth”) were the main contents; media (such as “reading” and “news”) was also considered. These topics were included because (1) information literacy is the basic literacy that citizens need to adapt to the information age, so it is highly necessary to attach importance to its cultivation in students, and, (2) with the development of education and health informatization, it is highly important to improve the information literacy of related users in these fields. Nevertheless, the breadth and depth of the research field must be strengthened.

The research on education and health information literacy is as rich as ever, especially during the COVID-19 pandemic (Wang et al., 2020; Wu et al., 2020b). In addition, more attention has been focused on the development of information literacy in various, diverse and specific fields in recent years. Liaqat et al. (2020) proved that financial information literacy can help manage earnings, Ahmad et al. (2020) investigated the relationship between CEO information literacy and innovation in enterprises, and Chebet and Cheruiyot (2020) identified the contributions of farmers’ information literacy to productivity and profitability. The topic is of great significance to improve citizens’ lives and work, promote social progress and eliminate the information gap.

In addition to enriching the content, studies on information literacy can increase the depth of research as follows. First, they focus on the role of information literacy in selected issues, such as the identification of fake news. Hagggar (2020) analyzed George Orwell’s diaries through an information literacy lens, which demonstrated how concerns regarding the need to evaluate information sources were represented. Second, we explore the application of information literacy in various scenarios. Pinto et al. (2020) used scale validation techniques and an exploratory factorial analysis to analyze the use of mobile technologies in the teaching–learning of information competencies. Third, we explore the influence mechanism of information literacy via new methods. Wu et al. (2020c) examined the impacts of parents on adolescents’ information literacy through a person-centered approach that utilized a latent profile analysis.

Innovation of evaluation methods based on data

The evaluation of information literacy is a highly important issue for improvement, while few studies have been conducted on “evaluation”. The share of the global topic “evaluation” is only 4.9%. From the perspective of the temporal dimension, information literacy evaluation initially focused on “self-efficacy”; gradually developed to “program”, “intervention” and experiment-related words; and has focused on “test” and “group” in recent years. From the perspective of subjects, NU emphasized the measurement of information literacy earlier, but it focused mainly on “self-efficacy”, which was absorbed by many subjects, especially for children’s use. Overall, the information literacy evaluation studies that were conducted were relatively traditional and lacked deep-seated evaluation and exploration.

In the evaluation of information literacy, typical standards and frameworks have been established (AASL&AECT, 1998; ACRL, 2000, 2015; IEA, 2013; ILFA, 2006; SCONUL,

1999, 2011; UNESCO, 2010, 2011, 2013). The current authoritative standards are used mainly for learning and education, which affirms the importance of information literacy for technology development and adaptation but shows a relative lack of situational specificity, which is consistent with the above discussion. However, scholars have expanded and constructed special information literacy evaluation indicators for various fields and objects (Niemelä et al., 2012; Zhou et al., 2020; Zhu et al., 2019), which provide satisfactory support for the precise evaluation of information literacy in different scenarios.

Regarding the methods of evaluating information literacy, most scholars use only self-designed questionnaires that consist of self-assessments with closed-ended test questions (Pinto et al., 2019). Other scholars have combined interviews (Walters et al., 2020), experiments (Ding & Ma, 2013) and other survey methods to evaluate the developmental level of information literacy. However, these methods lack enthusiasm and flexibility, and the data collection process requires substantial cooperation from users, which is time-consuming and laborious. With the development of the Internet, big data, artificial intelligence and other emerging technologies, massive and diverse process data can be recorded, which provides the possibility of data-driven evaluation. Various studies on technology behavior have used user process data; for example, Han et al. (2019) assessed the degrees and features of teachers' online participation in BL implementation, and Walters et al. (2020) used students' comments on a library instruction session to evaluate their information literacy. Kim et al. (2020) used online search behavior to identify differences between self-perceived eHealth literacy and performance in judging the authenticity of cancer information. These studies provide a new approach for process evaluation, but few data-driven evaluations of information literacy have been conducted using multispatial fusion data.

Conclusion

This study conducted a topic model dynamic analysis of the articles on information literacy research in the WOS core collection database from 2005 to 2019. The global topic and its popularity, topic similarity and correlation, temporal evolution of local topics, and subject diffusion of local topics were analyzed and presented.

For the first research question, nine global topics were identified by the LDA model—namely, “learning and education”, “library service”, “new digital technology”, “teacher ICT”, “health information”, “internet use”, “medium literacy”, “evaluation”, and “computer skill”—which presented the following characteristics: “learning and education” and “library service” were relatively stable and were covered by many papers. “New digital technology” and “teacher ICT” showed increasing trends, especially in EER, indicating that information literacy plays an important role in the process of education modernization. The research content of information literacy in the CS field was wider ranging. The development of technology is closely related to the requirements of information literacy, and it is the necessary result of the research and development of data-driven information literacy teaching and evaluation (Gómez-García et al., 2020). In the future, researchers in all fields should pay attention to the development of information literacy in the computer field and find a more suitable entry point through cross fusion.

For the second and third research questions, the words showed different temporal and subject characteristics in the local topics and focused on ability, technology, field, people, place and application in terms of information literacy. The temporal local topics of information literacy showed four evolutionary mechanisms: transferring, suspending,

crossing and growing. Crossing, which typically occurs on topics with widely used words, was the main mechanism; hence, the core topic words remained stable with few emerging topics. For the subject local topics of information literacy, by examining the diffusion of topics between each pair of subjects, five diffusion mechanisms were identified: absorbing, absorbing with division, absorbing with merging, paralleling, and interlacing. Among them, absorbing with division and absorbing were the main mechanisms, which supported the diffusion progress of information literacy studies. Various interdisciplinary subjects—such as PS, CS and ISIS—showed more interactivity, while the subjects highlighted their research preferences. In addition, information on the assessment of information literacy and related abilities in the field of health were disseminated. However, the current research on information literacy remains insufficient, few new research topics have emerged, a cooperation mechanism between disciplines has not been formed, and successful cases with universal practice are lacking. Many scholars have advocated the multidisciplinary integration of information literacy research and application in the early stage (Aharony, 2010; Weiner, 2011). Although research topics in different subjects have crossed, integration has not truly been achieved in terms of theoretical basis, research methods, practical exploration, etc., and more in-depth exploration is needed to form a new research paradigm (Huang et al., 2020).

Furthermore, it discussed the future research direction of information literacy to provide references for educators, researchers and practitioners. First, the rapid development and diversity of technology gives us higher requirements, which means that information literacy education should be strengthened, through such measures as training professional teachers and covering more audiences. (Huang et al., 2020). In addition, the research content of the impact of technology on information literacy should receive more attention. Second, the scope of information literacy will continue to expand and develop in more fields. In addition, the occurrence of the 2019 novel coronavirus (2019-nCoV) highlights the importance and urgency of improving information literacy for large-scale and long-term online teaching (Wu et al., 2020b) and identifying false public health information (Wang et al., 2020). Third, a gradual turn to data-driven information literacy evaluation research is needed because information behavior and information literacy are correlated (Hepworth et al., 2014). A large amount of user data is recorded in the information and intelligent environment, which can be used to conduct more real, accurate, multidimensional, undisturbed and continuous evaluations.

Acknowledgements This research has been made possible through the financial support of the National Natural Science Foundation of China under Grant Nos. 71904057 and 71871102, and the China Postdoctoral Science Foundation under Grant No. 2020M682453.

References

- AASL&AECT. (1998). Information power:Building partnerships for learning. In *Chicago:American Library Association Editions*.
- ACRL. (2000). Information literacy competency standards for higher education. Retrieved from: <http://www.ala.org/acrl/sites/ala.org/acrl/files/content/standards/standards.pdf>. Accessed 26 Oct 2019.
- ACRL. (2015). Framework for information literacy for higher education. Retrieved from: http://www.ala.org/acrl/standards/ils_recomm.pdf. Accessed 26 Oct 2019.
- Aharony, N. (2010). Information literacy in the professional literature: An exploratory analysis. *Aslib Proceedings: New Information Perspectives*, 62(3), 261–282. <https://doi.org/10.1108/00012531011046907>

- Ahmad, F., Widén, G., & Huvila, I. (2020). The impact of workplace information literacy on organizational innovation: An empirical study. *International Journal of Information Management*, 51(4), 1–12. <https://doi.org/10.1016/j.ijinfomgt.2019.102041>
- Aljanabi, A. R. A., & AL-Hadban, W. K. H. M. (2018). Information literacy and learners' satisfaction: the mediating role of social networking technology usage. *New Library World*, 119(11), 618–634. <https://doi.org/10.1108/ILS-08-2018-0091>
- American Library Association. (1989). *Presidential committee on information literacy*. American Library Association.
- Bawden, D. (2001). Information and digital literacies: A review of concepts. *Journal of Documentation*, 57(2), 218–259. <https://doi.org/10.1108/EUM0000000007083>
- Behrens, S. J. (1994). A conceptual analysis and historical overview of information literacy. *College and Research Libraries*, 55(4), 309–322. https://doi.org/10.5860/crl.55_04_309
- Bhardwaj, R. K. (2017). Information literacy literature in the social sciences and humanities: A bibliometric study. *Information and Learning Science*, 118(1–2), 67–89. <https://doi.org/10.1108/ILS-09-2016-0068>
- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python*. O'Reilly Media Inc.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022. <https://doi.org/10.1016/b978-0-12-411519-4.00006-9>
- Bruce, C. (2000). Information literacy programs and research: An international review. *Australian Library Journal*, 49(3), 209–218. <https://doi.org/10.1080/00049670.2000.10755921>
- Califf, C. B., & Brooks, S. (2020). An empirical study of techno-stressors, literacy facilitation, burnout, and turnover intention as experienced by K-12 teachers. *Computers and Education*, 157, 103971. <https://doi.org/10.1016/j.compedu.2020.103971>
- Celep, C., & Tülübaş, T. (2014). Effect of principals' technological leadership on teachers' attitude towards the use of educational technologies. *IFIP Advances in Information and Communication Technology*. https://doi.org/10.1007/978-3-662-45770-2_21
- Chebet, N., & Cheruiyot, J. K. (2020). Farmers' information literacy and productivity performance of smallholder horticulture in a highland zone, Kenya. *Journal of Scientific Research and Reports*, 26(6), 89–99. <https://doi.org/10.9734/jsrr/2020/v26i630274>
- Chen, B., Tsutsui, S., Ding, Y., & Ma, F. (2017). Understanding the topic evolution in a scientific domain: An exploratory study for the field of information retrieval. *Journal of Informetrics*, 11(4), 1175–1189. <https://doi.org/10.1016/j.joi.2017.10.003>
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *NAACL HLT 2019–2019 conference of the North American chapter of the association for computational linguistics: Human language technologies—proceedings of the conference* (pp. 1–14).
- Ding, R., & Ma, F. (2013). Assessment of university student web searching competency by a task-based online test: A case study at Wuhan University, China. *Electronic Library*, 31(3), 359–375. <https://doi.org/10.1108/EL-03-2011-0044>
- Farjon, D., Smits, A., & Voogt, J. (2019). Technology integration of pre-service teachers explained by attitudes and beliefs, competency, access, and experience. *Computers and Education*, 130, 81–93. <https://doi.org/10.1016/j.compedu.2018.11.010>
- Gómez-García, G., Hinojo-Lucena, F. J., Cáceres-Reche, M. P., & Navas-Parejo, M. R. (2020). The contribution of the flipped classroom method to the development of information literacy: A systematic review. *Sustainability*, 12(7), 1. <https://doi.org/10.3390/su12187273>
- Haggar, E. (2020). Fighting fake news: exploring George Orwell's relationship to information literacy. *Journal of Documentation*. <https://doi.org/10.1108/JD-11-2019-0223>
- Han, X., Wang, Y., & Jiang, L. (2019). Towards a framework for an institution-wide quantitative assessment of teachers' online participation in blended learning implementation. *The Internet and Higher Education*, 42, 1–12. <https://doi.org/10.1016/j.iheduc.2019.03.003>
- Haruna, H., & Hu, X. (2018). International trends in designing electronic health information literacy for health sciences students: A systematic review of the literature. *Journal of Academic Librarianship*, 44(2), 300–312. <https://doi.org/10.1016/j.acalib.2017.12.004>
- Hepworth, M., Almeahadi, F., & Maynard, S. (2014). A reflection on the relationship between the study of people's information behaviour and information literacy: Changes in epistemology and focus. *Library and Information Science*, 9, 51–66. <https://doi.org/10.1108/S1876-056220140000010018>
- Hsieh, P. N., Chuang, T. M., & Wang, M. L. (2013). A bibliometric analysis of the theses and dissertations on information literacy published in the United States and Taiwan. In: R. Howlett, L. C. Jain (eds.) *Smart innovation, systems and technologies* (vol. 1, pp. 337–348). Springer. https://doi.org/https://doi.org/10.1007/978-3-642-35452-6_35

- Huang, R., Wu, D., Kim, J., & Leung, B. T. H. (2020). Data and information literacy education: Methods, models, and challenges. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*. <https://doi.org/10.1145/3383583.3398546>
- Hughes, J. E., Cheah, Y. H., Shi, Y., & Hsiao, K. H. (2020). Preservice and inservice teachers' pedagogical reasoning underlying their most-valued technology-supported instructional activities. *Journal of Computer Assisted Learning*, 36(4), 549–568. <https://doi.org/10.1111/jcal.12425>
- IEA. (2013). Preparing for life in a digital age: The IEA international computer and information literacy study international report. Retrieved from: http://www.iea.nl/fileadmin/user_upload/Publications/Electronic_versions/ICILS_2013_International_Report.pdf. Accessed 26 Oct 2019.
- ILFA. (2006). Guidelines on information literacy for lifelong learning. Retrieved from: <http://www.ifla.org/files/assets/information-literacy/publications/ifla-guidelines-en.pdf>. Accessed 26 Oct 2019.
- Ketikidis, P., Dimitrovski, T., Lazuras, L., & Bath, P. A. (2012). Acceptance of health information technology in health professionals: An application of the revised technology acceptance model. *Health Informatics Journal*, 18(2), 124–134. <https://doi.org/10.1177/1460458211435425>
- Kim, S., Park, K., & Jo, H. S. (2020). Gap between perceived eHealth literacy and ability to use online cancer-related information. *Journal of Korean Medical Science*, 35(24), e187. <https://doi.org/10.3346/jkms.2020.35.e187>
- Kolle, S. R. (2017). Global research on information literacy: A bibliometric analysis from 2005 to 2014. *Electronic Library*, 35(2), 283–298. <https://doi.org/10.1108/EL-08-2015-0160>
- Kuhn, K. D. (2018). Using structural topic modeling to identify latent topics and trends in aviation incident reports. *Transportation Research Part C: Emerging Technologies*, 87, 105–122. <https://doi.org/10.1016/j.trc.2017.12.018>
- Kumar, K. (2014). A Scientometric study of digital literacy in online library information science and technology abstracts (LISTA). *Library Philosophy and Practice*, 1, 1–13.
- Li, D., Mei, H., Shen, Y., Su, S., Zhang, W., Wang, J., Zu, M., & Chen, W. (2018). ECharts: A declarative framework for rapid construction of web-based visualization. *Visual Informatics*, 2, 136–146. <https://doi.org/10.1016/j.visinf.2018.04.011>
- Li, F., Li, M., Guan, P., Ma, S., & Cui, L. (2015). Mapping publication trends and identifying hot spots of research on Internet health information seeking behavior: A quantitative and co-word biclustering analysis. *Journal of Medical Internet Research*, 17(3), e81. <https://doi.org/10.2196/jmir.3326>
- Liaquat, F., Mahmood, K., & Ali, F. H. (2020). Demographic and socio-economic differences in financial information literacy among university students. *Information Development*, 2, 1–12. <https://doi.org/10.1177/0266666920939601>
- Limberg, L., Sundin, O., & Talja, S. (2012). Three theoretical perspectives on information literacy. *Human IT*, 11(2), 93–130.
- Lloyd, A. (2005). Information literacy: Different contexts, different concepts, different truths? *Journal of Librarianship and Information Science*, 37(2), 82–88. <https://doi.org/10.1177/0961000605055355>
- Marcum, J. W. (2002). Rethinking information literacy. *The Library Quarterly*, 72(1), 1–26. <https://doi.org/10.1086/603335>
- Martzoukou, K., & Sayyad Abdi, E. (2017). Towards an everyday life information literacy mind-set: A review of literature. *Journal of Documentation*, 73(4), 634–665. <https://doi.org/10.1108/JD-07-2016-0094>
- Nazim, M., & Ahmad, M. (2007). Research trends in information literacy: A bibliometric study. *Journal of Information Management*, 49:53–62. <https://doi.org/https://doi.org/10.17821/SRELS/2012/V49I2/44089>
- Niemelä, R., Ek, S., Eriksson-Backa, K., & Huotari, M. L. (2012). A screening tool for assessing everyday health information literacy. *Libri*, 62(2), 125–134. <https://doi.org/10.1515/libri-2012-0009>
- Panda, I., Maharana, B., & Chhatar, D. C. (2013). The journal of information literacy: A bibliometric study. *International Journal of Scientific and Research Publications*, 3(3), 1–7.
- Park, M.-K., & Kim, H.-J. (2011). A bibliometric analysis of the literature on information literacy. *Journal of the Korean Society for Information Management*, 28(2), 53–63. <https://doi.org/10.3743/kosim.2011.28.2.053>
- Pinto, M. (2015). Viewing and exploring the subject area of information literacy assessment in higher education (2000–2011). *Scientometrics*, 102(1), 227–245. <https://doi.org/10.1007/s11192-014-1440-2>
- Pinto, M., Caballero, D., Sales, D., & Fernández-Pascual, R. (2020). MOBILE-APPS questionnaire: Developing and validating a scale to measure the attitudes and perceptions of undergraduate students on mobile information literacy. *Journal of Librarianship and Information Science*, 52(2), 096100062090226. <https://doi.org/10.1177/0961000620902260>

- Pinto, M., Córdón, J. A., & Díaz, R. G. (2010). Thirty years of information literacy (1977–2007): A terminological, conceptual and statistical analysis. *Journal of Librarianship and Information Science*, 42(1), 3–19. <https://doi.org/10.1177/0961000609345091>
- Pinto, M., Escalona-Fernández, M. I., & Pulgarín, A. (2013). Information literacy in social sciences and health sciences: A bibliometric study (1974–2011). *Scientometrics*, 95(3), 1071–1094. <https://doi.org/10.1007/s11192-012-0899-y>
- Pinto, M., Fernández-Pascual, R., Caballero-Mariscal, D., & Sales, D. (2020). information literacy trends in higher education (2006–2019): Visualizing the emerging field of mobile information literacy. *Scientometrics*, 124(2), 1479–1510.
- Pinto, M., Fernández-Pascual, R., Caballero-Mariscal, D., Sales, D., Guerrero, D., & Uribe, A. (2019). Scientific production on mobile information literacy in higher education: A bibliometric analysis (2006–2017). *Scientometrics*, 120(1), 57–85. <https://doi.org/10.1007/s11192-019-03115-x>
- Pinto, M., Isabel Escalona, M., Pulgarín, A., & Uribe-Tirado, A. (2014). The scientific production of Ibero-American authors on information literacy (1985–2013). *Scientometrics*, 102(2), 1555–1576. <https://doi.org/10.1007/s11192-014-1498-x>
- Pinto, M., Pulgarín, A., & Escalona, M. I. (2014). Viewing information literacy concepts: A comparison of two branches of knowledge. *Scientometrics*, 98(3), 2311–2329. <https://doi.org/10.1007/s11192-013-1166-6>
- Pinto, M., Sales, D., & Fernández-Pascual, R. (2019). Gender perspective on information literacy: An interdisciplinary and multidimensional analysis within higher education settings. *Library & Information Science Research*. <https://doi.org/10.1016/j.lisr.2019.100979>
- Prado, J. C., & Marzal, M. A. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*, 63(2), 123–134. <https://doi.org/10.1515/libri-2013-0010>
- Rehurek, R., & Sojka, P. (2010). Software framework for topic modelling with large corpora. In *Proceedings of the LREC 2010 workshop on new challenges for NLP frameworks* (pp. 45–50).
- Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, 128, 13–35. <https://doi.org/10.1016/j.compedu.2018.09.009>
- SCONUL Working Group on Information Literacy. (1999). *Information skills in higher education: A SCONUL position paper*. Society of College, National and University Libraries.
- SCONUL Working Group on Information Literacy. (2011). *The SCONUL seven pillars of information literacy: core model for higher education*. Society of College, National and University Libraries.
- Sievert, C., & Shirley, K. (2014). LDAvis: A method for visualizing and interpreting topics. *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*. <https://doi.org/10.3115/v1/w14-3110>
- Sproles, C., Detmering, R., & Johnson, A. M. (2013). Trends in the literature on library instruction and information literacy, 2001–2010. *Reference Services Review*, 41(3), 395–412. <https://doi.org/10.1108/RSR-03-2013-0014>
- Stevens, K., Kegelmeyer, P., Andrzejewski, D., & Buttlar, D. (2012). Exploring topic coherence over many models and many topics. In *EMNLP-CoNLL 2012–2012 joint conference on empirical methods in natural language processing and computational natural language learning* (pp. 952–961).
- Stopar, K., & Bartol, T. (2019). Digital competences, computer skills and information literacy in secondary education: Mapping and visualization of trends and concepts. *Scientometrics*, 118(2), 479–498. <https://doi.org/10.1007/s11192-018-2990-5>
- Swamy, N. (2016). *Cluster purity visualizer*. <https://Gist.Github.Com/Nswamy14/E28ec2c438e9e8bd>. Accessed 26 Oct 2019.
- Tallosli, S. B., & Mulla, K. R. (2016). A bibliometric analysis of journal of information literacy (2011 to 2015). *ISST Journal of Advances in Librarianship*, 7(2), 44–51.
- Taşkın, Z., Doğan, G., & Şencan, I. (2013). Analyzing the intellectual structure of world information literacy literature through citations and co-citations. *Communications in Computer and Information Science*, 397, 54–60. https://doi.org/10.1007/978-3-319-03919-0_6
- Tuominen, K., Savolainen, R., & Talja, S. (2005). Information literacy as a sociotechnical practice. *Library Quarterly: Information, Community, Policy*, 75(3), 329–345. <https://doi.org/10.1086/497311>
- UNESCO. (2010). Towards media and information literacy indicators. Retrieved from: <https://www.ifla.org/files/assets/information-literacy/publications/towards-media-and-information-literacy-indicators.pdf>. Accessed 26 Oct 2019.
- UNESCO. (2011). UNESCO ICT competency framework for teachers. Retrieved from: <http://unesdoc.unesco.org/images/0021/002134/213475e.pdf>. Accessed 26 Oct 2019.
- UNESCO. (2013). Global media and information literacy assessment framework. Retrieved from: <http://unesdoc.unesco.org/images/0022/002246/224655e.pdf>. Accessed 26 Oct 2019.

- Wai-yi, B. C. (1998). An information seeking and using process model in the workplace: A constructivist approach. *Asian Libraries*, 7(12), 375–390. <https://doi.org/10.1108/10176749810241919>
- Walters, W. H., Sheehan, S. E., Handfield, A. E., López-Fitzsimmons, B. M., Markgren, S., & Paradise, L. (2020). A multi-method information literacy assessment program: Foundation and early results. *Portal: Libraries & the Academy*, 20(1), 101. <https://doi.org/10.1353/pla.2020.0006>
- Wang, F., An, L., Huang, R., & Wang, X. (2020). Scientific response and reflection of public health emergencies: An interview with LIS experts on COVID-19. *Documentation, Information & Knowledge*, 2, 1–14.
- Weiner, S. (2011). How information literacy becomes policy: An analysis using the multiple streams framework. *Library Trends*, 60(2), 297–311. <https://doi.org/10.1353/lib.2011.0037>
- Wu, Dan, Li, X., Xu, S., Dong, J., Fan, S., & Gui, D. (2020). Research on the usage and development of information literacy theory in the recent ten years. *Library Journal*, 39(1), 26–35. <https://doi.org/https://doi.org/10.13663/j.cnki.lj.2020.01.002>.
- Wu, Di., Yu, L., Rao, J., Zhou, C., & Chen, M. (2020). Challenges and promotion strategies of information literacy of teachers and students for large-scale and long-term online teaching. *E-Education Research*, 41(5), 12–17.
- Wu, Di., Yu, L., Yang, H. H., Zhu, S., & Tsai, C. C. (2020). Parents' profiles concerning ICT proficiency and their relation to adolescents' information literacy: a latent profile analysis approach. *British Journal of Educational Technology*, 51(6), 2268–2285. <https://doi.org/10.1111/bjet.12899>
- Xie, Q., Zhang, X., Ding, Y., & Song, M. (2020). Monolingual and multilingual topic analysis using LDA and BERT embeddings. *Journal of Informetrics*, 14(3), 101055. <https://doi.org/10.1016/j.joi.2020.101055>
- Zhou, C., Wang, H., Liang, X., & Chen, M. (2020). Development and application of information literacy assessment tool for primary and secondary school teachers. In *Proceedings of the 2020 ninth international conference of educational innovation through technology* (pp. 70–75).
- Zhu, S., Wang, Y., Wu, D., Shi, Y., & Yang, H. H. (2019). Development and validation of information literacy assessment tool for primary students. *International Symposium on Educational Technology (ISET)*, 2019, 7–11.
- Zurkowski, P. G. (1974). The information service environment relationships and priorities. *National Commission on Libraries and Information Science*.