International Journal of Information Technologies and Systems Approach
Volume 16 • Issue 2

An Intelligent Retrieval Algorithm for Digital Literature Promotion Information Based on TRS Information Retrieval

Tong Ni, Xi'an Innovation College of Yan'an University, Xi'an, China*

ABSTRACT

The world has entered the information age, and one of the main factors of social development is various information resources. Word processing is one of the first information technologies to be developed that has grown very rapidly and successfully. As a category of multimedia resources, documents are widely used in institutions, such as corporations, governments, and digital libraries. Traditional search technology has been unable to meet the needs of this development, and how to find information in this sea of digital information has become an urgent problem. Therefore, this study has conducted a research experiment on promoting digital literature for text retrieval system information retrieval. The experimental data have shown that 310 (73.81%) students wanted WeChat and Weibo as a way of digital literature retrieval and promotion. Moreover, 267 (63.57%) students wanted e-books as a way to promote, and 172 (40.95%) students would like to participate in novel digital literature retrieval promotion activities.

KEYWORDS

Digital Literature, Information Retrieval, Literature Promotion, Retrieval Algorithm, TRS Information Retrieval

1. INTRODUCTION

Digital library is developing rapidly in the era of information explosion and network technology. As a one-stop service for information sources scattered on the Internet, digital libraries are constantly expanding and improving in space, time and service functions. Digital library provides not only lending, reference, information retrieval (IR) and personalisation services but also customised specific services, such as document reproduction, document distribution, interlibrary loan, virtual consultation, video-on-demand and email personalisation using web technologies. However, library users believe that the rapid development of information technology has completely changed the way information is stored, transmitted and acquired whilst also shortening the process or cycle of using and creating knowledge and changing the way and channel for acquiring information. Knowledge consumers are often also knowledge producers, which leads to an exponential increase in the rate of knowledge or information creation. Therefore, people' most common perception in online environments is information overload, not information scarcity. In this study, the digital literature promotion information intelligent retrieval

DOI: 10.4018/ijitsa.318458 *Corresponding Author

algorithm is studied based on text retrieval system (TRS) IR to make some contributions to the effective management of literature resources.

Retrieval of digital document information currently faces two major problems: one is how to quickly determine the retrieval scope across multiple information resources, and the other is how to filter out the 'relevant information' that has been found. Based on these problems, users have proposed higher requirements for the library's intelligent IR service to query digital document data more accurately and quickly. Thus, digital resources are numerous, and library users are faced with considerable information and are unable to obtain the resources they need timely. In reality, the needs of library users are complex and diverse. Users cannot obtain access to all the information provided by the library, and the needs of individual users are only a part of it. Thus, the need for IR capabilities is becoming increasingly urgent.

The innovations of this study are as follows: (1) Digital document retrieval based on the TRS IR algorithm is analysed, and a digital document promotion method based on TRS IR is proposed; (2) The experiment of digital document retrieval is carried out.

2. RELATED WORK

This study examines some techniques of digital literature, which can be fully applied to research in this field. Cunha and Frota (2017) aimed to advance the field of telematics research in literary texts. Ahmad and Zhu (2019) studied digital literature in the Middle East and its impact on politics. Rebecca and Kalamani (2020) aimed to facilitate the exchange of ideas and increase easy access to literature through e-books, blogs and web pages (including audio and graphics). Onuoha (2018) aimed to examine the digital literature in the diaspora: role switching and self-construction in Iheloja's autobiography. Edmonds (2019) believed that the use of digital technology to teach literature assignments is a collection of assignments that utilise technology in college-level literature classes. Ubiparipovi, Matkovi and Mari (2020) aimed to identify key success factors for purposeful digital transformation through a literature review. These methods have provided some references for research. However, owing to the short time and small sample size of the relevant research, they have not been recognised by the public.

Based on IR, the following related materials were reviewed to optimise digital literature research. Na and Kim (2017) believed that document length normalisation is one of the basic components in retrieval models. Chen, Tsutsui and Ding (2017) used datasets from IR publications to analyse topic trends, evolution dynamics and semantic word movement in the domain. Kalyanasundaram, Abraham and Ramachandran (2017) evaluated the impact of mind-mapping technology on IR of medical students. Gianmaria, Georgeta and Nicola (2017) discussed how data could be semantically annotated and linked to enhance their interpretation, sharing and reuse. Han, Chen and Tian (2018) argued that query expansion in IR systems requires knowledge bases to consider semantic relationships between words. These methods have provided a sufficient literature basis for the study of digital literature promotion of TRS IR.

3. OVERVIEW OF TRS IR AND DIGITAL LITERATURE RETRIEVAL

The Internet, also known as the international network, refers to a large network connected by networks. These networks are connected by a set of universal protocols to form a single, large, logical international network. With the rapid development of information technology, the Internet is spreading worldwide, and the speed of collecting and disseminating documents has reached an unprecedented level. Therefore, the current digital document retrieval technology is increasingly unable to meet the needs of the public. Therefore, this study aims to take the development of intelligent retrieval of digital documents to a higher level by studying TRS IR.

3.1 Overview of TRS IR

TRS focuses on Chinese data processing and search technology research. Text retrieval, also known as natural language retrieval, refers to a system that searches for words in natural language directly through computers without any indexing of documents. TRS full-text database system is most commonly used in media, education, business, government, scientific research and other fields and is the most stable, mature and reliable IR product in China (Taniguchi, 2018). TRS application development interface could exchange information with the TRS server. Figure 1 shows the structure of the TRS full-text database system.

The TRS structure has basic functions, such as search, index, web-based development interface, user-friendly interface and secondary application development interface. It also has its own comprehensive management of unstructured data, powerful application support architecture, efficient real-time intelligent full-text search, reliability, high security, maturity and openness.

TRS IR technology focuses on analysing the efficiency and accuracy of massive IR in China, including technical research and application practices in various fields, such as dictionary technology, word extraction technology, indexing technology and relevance ranking of retrieval results (Koerber, 2020).

IR has broad and narrow senses. Broadly speaking, IR means to store information and organise in a certain way and get relevant information according to the needs of users. Information search in a narrow sense means only extracting basic information from data collection, and it only involves that part of IR in a broad sense that corresponds to what is commonly called information extraction. Moreover, information can be classified into images, text, numbers and speech, based on how the information is presented. Although information can be presented in many ways, the text is the most commonly used, which is the most concise and abstract way of presenting information.

Currently, the search for information has evolved, becoming networked and intelligent. Adapting to the needs of network intelligence and personalisation is now a new trend in the development of information technology (Kamal, 2019). Figure 2 shows the conceptual model of IR.

IR is a process. The factor that determines whether a file is indexed is its similarity to the content of the file retrieved by the user, and it is quantified by computing the expressive features of the content.

IR techniques help people find useful information from existing information, saving much time for them. Searching for information on the web means searching for information published on the Internet, and the most important search tool today is the search engine. The core of an IR system is a

Figure 1. TRS full-text database system structure

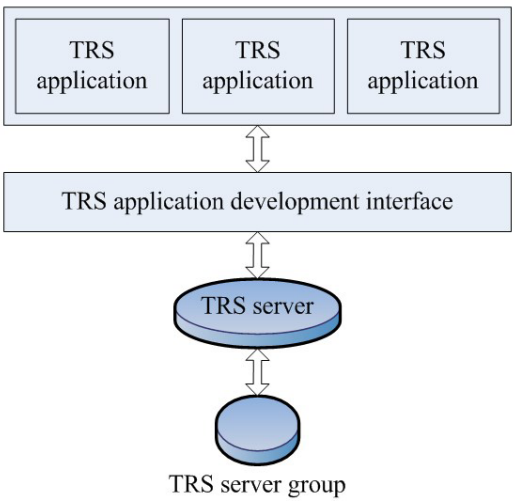
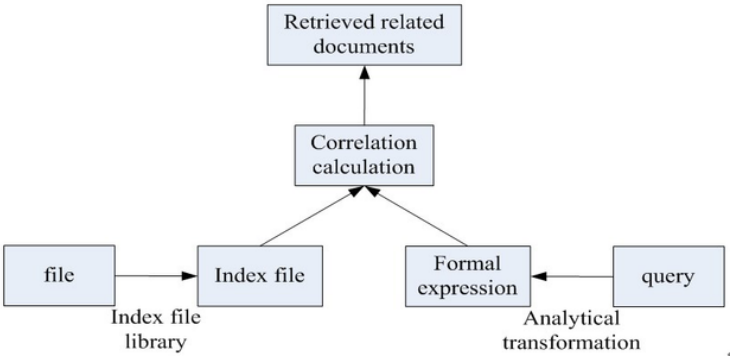


Figure 2. Conceptual model of IR

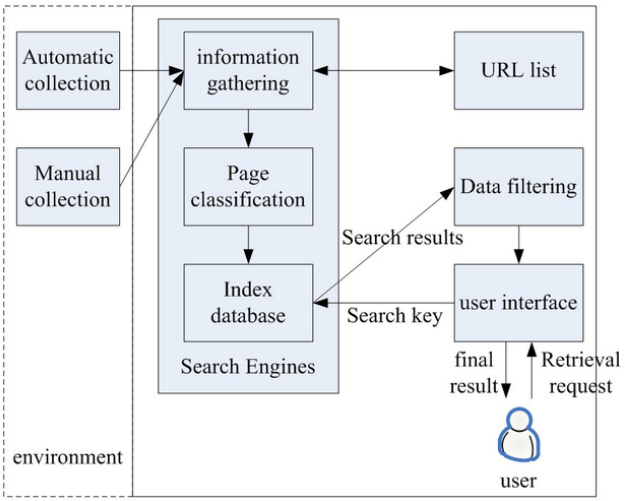


search engine, which resides on a search server and performs search tasks (Hobbs & Viinalass-Smith, 2018). Figure 3 depicts the basic structure of a typical IR system.

The workflow of the IR system is as follows: The information collection module of the search engine needs to collect information resources from the network environment, including manual and IR tools. Once the information collection module completes the task of collecting information, it sends the collected information sources back to the search engine, sorts and indexes these information sources, and then stores them in the created index database. The search engine provides users with an interface for searching for information on the Internet, and users use this interface to query and search. After submitting a search query, the search engine displays the final search results to the user in the form of hypertext links through which the user can access relevant information resources.

The classification of IR systems is mainly determined by the classification of search engines. The current search engines are mainly divided into directory search engines, which classify the collected information according to the directory; IR program-based search, which automatically collects and indexes data by IR programs; the client-side search engine, which uses server-side search software to search files. To find more files, the hyperlinks in the files need to be used until the requirements

Figure 3. Basic structure of a typical IR system



are met. Distributed search engines, which exchange intermediate data between index servers and proxy queries (Cabanac, Frommholz, & Mayr, 2018).

Information extraction techniques and methods can be divided into two categories: methods based on statistical and semantic data. Bayesian, Boolean, extended Boolean and vector spaces are statistical models for information extraction. Extended Boolean and generalised-based combinations are vector space models.

The classic algorithm for calculating the weight of index words is the term frequency and document turnover frequency (TF*IDF) method. The frequency with which an index term appears in a document is an indication of its importance in the document. Therefore, a simple way to determine the weight of an index word is to use the frequency of the word in the document to represent the weight. The inverse proportion of the total number of documents in which the word appears is the weight of the search word. In the formula, *y* is the frequency. The formula is as follows:

$$w_{i,j} \infty IDF_i = log\left(\frac{y}{y_i}\right) \tag{1}$$

Amongst them, the total number of documents in the datasets is represented by *y*, and the total number of documents with index words is represented by y_i, which reflects the general statistical characteristics of the importance of index words in the complete datasets. Therefore, the weights of the indexed terms in the document are calculated as follows:

$$w_{i,j} = TF_{i,j} \cdot IDF_i = num_i \cdot log\left(\frac{y}{y_i}\right) \tag{2}$$

The above equation shows that words with high frequency in one document and low frequency in the whole dataset are more important (Chawla, 2017).

The most important improvement of TF is the adjustment of word frequency, which is represented as the interval size [0,1]. Thus, the total number of words in the file or the frequency of the most frequent word in the file is as follows:

$$TF_{i,j} = \frac{freq_{i,j}}{max_k\left\{freq_{k,j}\right\}} \tag{3}$$

$$w_{i,j} = TF_{i,j} \cdot IDF_i = \frac{freq_{i,j}}{max_k\left\{freq_{k,j}\right\}} \cdot log\left(\frac{y}{y_i}\right) \tag{4}$$

The log-frequency method does not apply regularisation factors, such as text length or maximum word frequency, and it minimises the effect of word rate on TF sampling:

$$TF_{i,j} = log\left(freq_{i,j}\right) + 1 \tag{5}$$

If the terms and documents queried by users are quantified according to Eq. (2), their similarity can be calculated. Let *d* be a document vector and *q* be a user query word vector; then, the conditions that must be met for the value of $Simarilaty\left(d,q\right)$ are as follows:

- Non-negativity:

$$Simarilaty\left(d,q\right) \tag{6}$$

- Symmetry:

$$Simarilaty\left(d,q\right)=Simarilaty\left(q,d\right) \tag{7}$$

The normalised $Simarilaty\left(d,q\right)$ value range is usually any value in the interval [0,1].

Traditional vector correlation algorithms include Euclidean distance, block distance and cosine similarity method, of which the cosine similarity method and distance-based method are the most typical (Djenouri, Belhadi, & Belkebir, 2018).

A simple way to calculate similarity is to directly calculate the inner product of two Y-dimensional vectors, as shown in Eq. (8):

$$Simarilaty\left(q,d\right)=\sum_{i=1}^{Y}w_{i,j}\cdot w_{i,q} \tag{8}$$

If the two Y-dimensional vectors are cosine-regular, the inner product is the cosine of the angle between the two vectors, and the cosine similarity is calculated as follows:

$$Simarilaty\left(q,d\right)=\frac{\sum_{i=1}^{Y}w_{i,j}\cdot w_{i,q}}{\sqrt{\sum_{i=1}^{Y}\left(w_{i,j}\right)^{2}}\cdot\sqrt{\sum_{i=1}^{Y}\left(w_{i,q}\right)^{2}}} \tag{9}$$

Similar to the cosine similarity method, the documents and user query conditions *q* in the complete datasets are represented in the corresponding vector form, which gives:

$$Simarilaty\left(q,d\right)=\frac{1}{L_{p}\left(q,d\right)+1} \tag{10}$$

$$L_{p}\left(q,d\right)=\left[\sum_{i}\left|q_{i}-d_{i}\right|^{p}\right]^{l/p} \tag{11}$$

When p = 1, this measure becomes a block distance; when p = 2, this measure becomes a Euclidean distance; then, when $p=\infty$, this measure becomes a maximum directed distance (Holzapfel, Sturm, & Coeckelbergh, 2018).

As the traditional vector space model handles the features of the vector space uniformly without considering the location and internal meaning of these features in the document, some drawbacks appear in the page processing. The researchers propose two major improvements to address this shortcoming: HTML and hyperlink improvements.

In the hyperlink correction, the algorithm first uses the TF*IDF rule to calculate the weight of the feature expression in the target document. Then, the weight of the feature expression of the linked L documents in the target document environment is added to obtain the corrected weight. The details are as follows:

$$w_{tk} = \frac{tf\left(t_k, p\right)}{\sum_{i=1}^{x} tf\left(t_i, p\right)} * \log \frac{Y_{web}}{df\left(t_k\right)} \tag{12}$$

Amongst them, x is the number of attribute elements contained in file p, the total number of files in the file set is Y_{web} and $df\left(t_k\right)$ is the number of files in the t_k -file set, that is:

$$w_{tk}^{L} = \frac{1}{Dim}\left[\sum_{i=1}^{L}\sum_{j=1}^{Y_i} \frac{w_{t_k}^{p_{ij}}}{d\left(w, w^{p_{ij}}\right)}\right] \tag{13}$$

Amongst them, $w_{t_k}^{p_{ij}}$ is the weight of the feature element in the j-th linked document in the i-th row.

3.2 Overview of Digital Literature Search

Documentation is an inexhaustible wealth of knowledge in human society. The international community attaches great importance to developing and utilising information resources in this field. With the advancement of Internet and database technology, the use of advanced technologies, such as data mining, data storage and knowledge management, to create various specialised social knowledge systems and data analysis systems now plays an important role in promoting scientific and economic development.

The popularity and continuous development of the Internet have changed the source and form of documents, which are no longer limited to printed documents, but also include electronic versions of various types of information on the Internet. This new type of literature resource has become the third largest source of information after journals and books. More and more countries realise that digital documents are the research in the information age and the basic information source needed for economic development and social civilisation progress.

Currently, document use in developed countries is organised on a social basis. After 2000, the United States began to establish and continuously expand electronic file archives one after another. It builds books, archives, news and other information into comprehensive social information, to establish a document archive that is generally distributed and centrally managed, socialised and shared.

Considering the current situation in China, the digitisation of literary resources needs to be sped up. Considering the economic and technical conditions of the library and the needs of the entire collection, libraries must allocate certain human, material and financial resources to digitise documents and take reasonable measures and decisions. Priority should be given to digitising and developing relevant, valuable, distinctive, digitally valuable and technologically advanced literature.

For libraries, the choice of digital objects should be purposeful. The appropriate scope of digitisation should be determined according to the tasks, nature, needs of users, actual collection

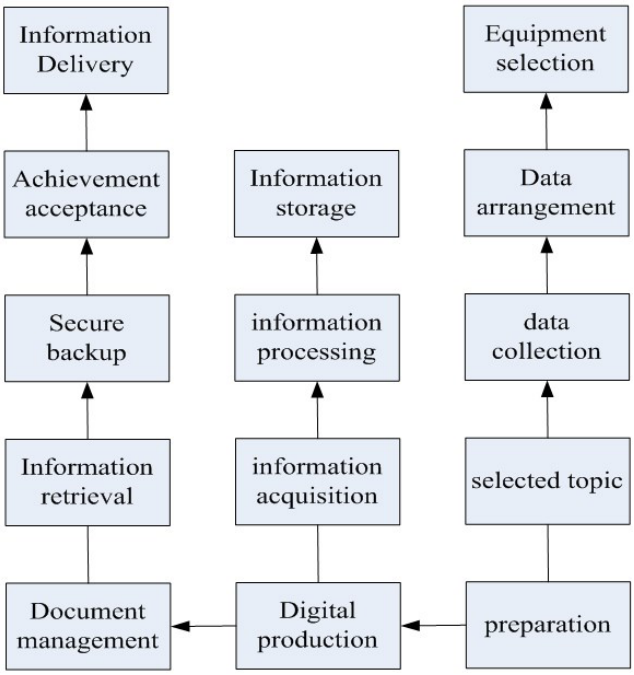
structure and development trends of schools. In particular, key literature disciplines should be selected in a targeted manner to meet the information needs of the majority of literature resource users as much as possible. Figure 4 shows the workflow of document digitisation.

The digitisation process of library documents and information resources is one of the most important stages in the entire digitisation project, which can be divided into three stages: the pre-digitisation stage, the production and processing of digital literature and information media and the publication and use of digital literature. The management and publication of digitised documents include securing data, extracting results, searching for information, publishing and providing information.

Documentation resources are provided in media, such as graphics, images, text, sound and video. Library information administrators need to choose the most appropriate digitisation method according to the characteristics and requirements of different media types. The digitisation methods that can be used for text documents include manual input, scanning input, optical character recognition and others. The digitisation methods of graphic literature resources include CAD redrawing, scanner scanning and a combination of the first two methods. (Computer-aided design software, which is used for two-dimensional drawing, detailed drawing, design documents and basic three-dimensional design, has become a widely popular drawing tool worldwide. AutoCAD has a good user interface, and you can perform various operations through interactive menus or command lines.) The digitisation of image files is to convert collected images, such as paintings and calligraphy works, microfilms and photographs, into digital image files suitable for computer processing through scanning and quantification. Digitisation of sound is the conversion of continuous analogue sound signals into discrete digital sound signals. It can be processed on a computer through operations such as sampling and quantisation by an analogue-to-digital converter. Digitisation of video files means that analogue video signals are captured, sampled and converted into digital data values that can be recognised and processed by a computer.

Advances in IT technology have made people use more and more physical objects to process content. The description of this object, that is, metadata, has become the core of IT systems, which is

Figure 4. Document digitisation workflow



different for different types of documents. However, in the process of use, transmission and sharing, a set of relatively large and fixed descriptive information can be uniformly formed, which is the metadata structure used in a unified document resource retrieval application.

According to the analysis of the heterogeneity and metadata characteristics of files, combined with the results of continuous analysis of file access logs, the retrieval and use of file resources are a comprehensive search based on heterogeneous resources. The retrieval tools and methods include secondary retrieval of retrieval results, advanced intelligent retrieval, simple general retrieval, advanced retrieval based on subject words and others, as well as voice search with humane care and location search with a higher level of specialisation. The search support functions include suggestions for related words, suggestions for search terms, statistics for classifying search results and so on. Combining the above-mentioned heterogeneity of resource attributes and document metadata with the analysis of the functions and performance requirements of document resource query, Figure 5 shows the application framework of TRS document resource retrieval.

Using the information management function, the storage of metadata and the management of comprehensive file resources are realised. Words are automatically extracted from the metadata of document resources based on word extraction dictionaries integrated with the platform using technologies such as word extraction, indexing and dictionaries. Thus, the search for Chinese content becomes more accurate and utilises information indexing for faster search results. The database server software is used to realise the search function of the unified resource search platform.

4. PROMOTION OF DIGITAL LITERATURE FOR IR

4.1 Status Quo of Digital Literature Retrieval

Based on IR technology, this section conducted a questionnaire analysis of the current status of digital document retrieval. It started from three aspects: the purpose of digital literature retrieval, the method of digital literature retrieval and the duration of digital literature retrieval. Questionnaires were randomly distributed to students in school. Online questionnaires were mainly distributed through email, Weibo and WeChat to avoid the geographical limitation of the research. There were 420 valid samples, 190 for boys and 230 for girls. The recovery rate of experimental data is not 100%, and the effective recovery rate is also 100%. Figure 6 shows the purpose of digital literature retrieval for students of different genders (multiple choice).

Figure 6(a) shows the distribution of search purposes for boys' digital literature, and Figure 6(b) shows the distribution of search goals for girls' digital literature. As shown in Figure 6, there were 52 boys and 68 girls who searched for the purpose of the professional study; 98 boys and 167 girls were for leisure and entertainment; 85 boys and 114 girls were for the purpose of expanding knowledge; 26 boys and 53 girls were for exams and homework; 18 boys and 26 girls were for other purposes. As

Figure 5. TRS document resource retrieval application framework

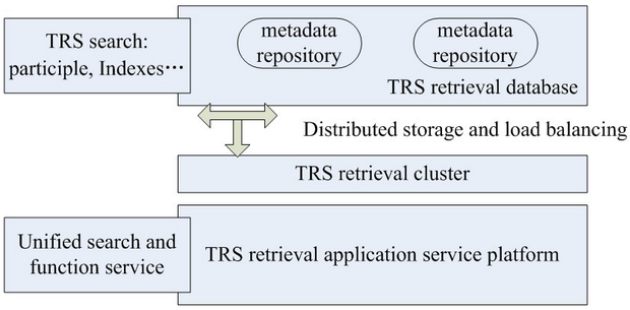
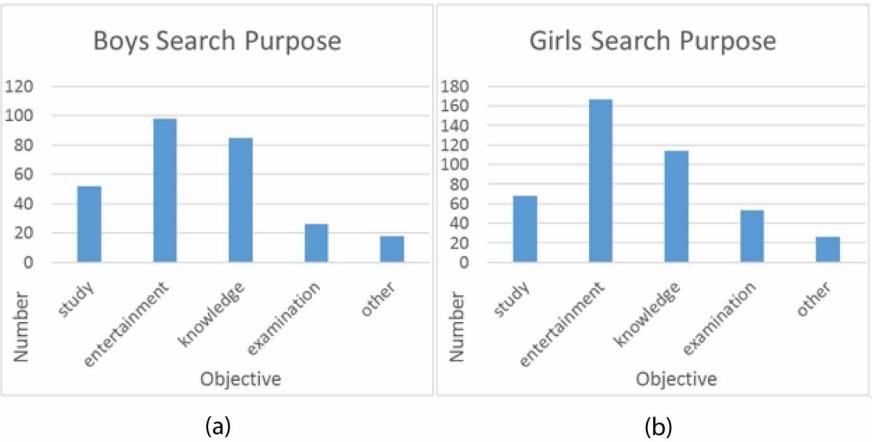


Figure 6. Purpose of digital literature retrieval for students of different genders (multiple choice)

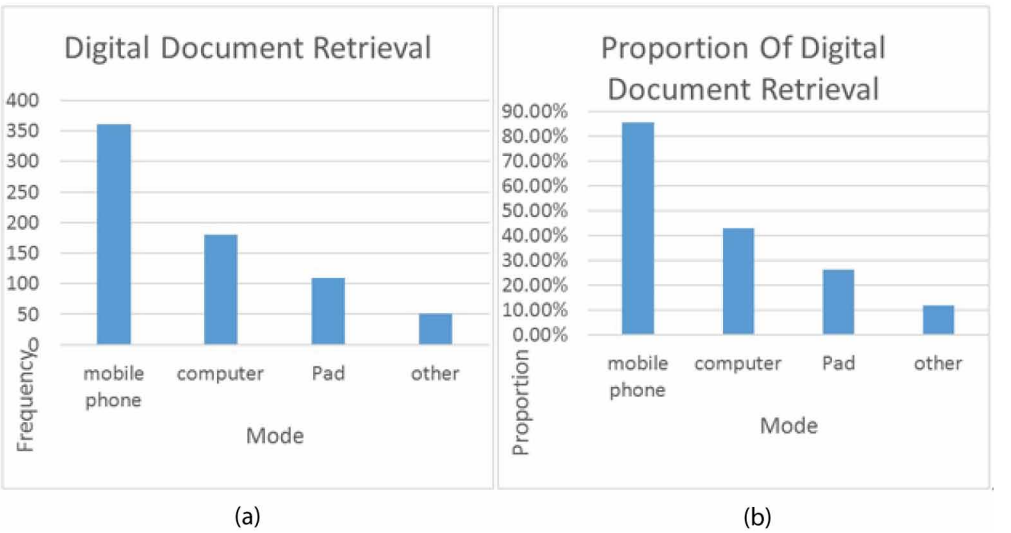


(a) (b)

for the purpose of digital literature retrieval, there is little difference between boys and girls. Figure 7 depicts the method of searching digital literature for students (multiple choice).

Figure 7(a) presents the distribution of students' digital literature retrieval methods, and Figure 7(b) shows the proportion of students' digital literature retrieval methods. As shown in Figure 7, 360 students used mobile phones for literature retrieval, accounting for 85.71%; 180 people used computers for literature retrieval, accounting for 42.86%; 110 people used a tablet for literature retrieval, accounting for 26.19%; there were 50 people used other methods for retrieval, accounting for 11.9%. From the above data, when students chose digital literature retrieval methods, they mainly used mobile phones, computers and tablets. The survey found that with the full popularity of smartphones, the proportion of current students owning smartphones was 100%. Figure 8 shows the duration of student digital literature retrieval.

Figure 7. Student's digital literature retrieval method (multiple choice)



(a) (b)

Figure 8. Time of digital literature retrieval of student

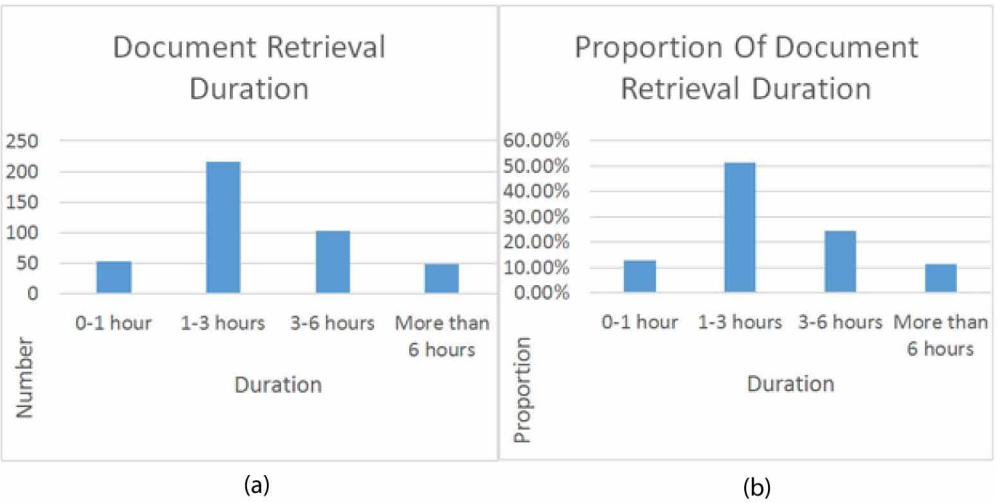


Figure 8(a) shows the time distribution of students' digital literature retrieval, and Figure 8(b) depicts the proportion of students' digital literature retrieval time. As shown in Figure 8, 53 students spent 0–1 hour searching the literature, accounting for 12.62%; 216 students spent 1–3 hours, accounting for 51.43%; 103 students spent 3–6 hours, accounting for 24.52%; and 48 people spent more than 6 hours, accounting for 11.43%. The above data showed that half of the students needed 1–3 hours to retrieve digital literature, and the retrieval time was longer than that of the public retrieval, indicating that students were a group with a high degree of popularity of national retrieval.

4.2 Promotion of Digital Literature Retrieval

This section combined the analysis of the digital literature retrieval in the previous part and conducted an in-depth investigation on the promotion of digital literature retrieval. Using the method in the previous part, it was analysed from three aspects: whether the school has a promotion, whether the school provides convenient channels and the expected promotion method. Table 1 shows the statistics on whether the school has carried out digital literature promotion activities.

Table 1 shows that 220 people, accounting for 52.38%, knew that the school had carried out promotion activities; 169 people, accounting for 40.24%, knew that the school had not carried out promotion activities; 31 people, accounting for 7.38%, did not know whether they had carried out promotion activities. From the above data, the proportion of schools with and without promotion activities was not much different, indicating that the promotion of digital literature retrieval in schools was far from enough. Table 2 shows the situation of whether the school provides convenient digital literature retrieval channels.

Table 1. Whether the school has carried out digital literature promotion activities

Option	Number	Proportion
Have	220	52.38%
No	169	40.24%
I don't know	31	7.38%

Table 2. Whether the school provides a convenient digital literature search channel

Option	Number	Proportion
Have	237	56.43%
No	158	37.62%
I don't know	25	5.95%

Table 2 shows that 237 people, accounting for 56.43%, said that the school provided a convenient channel; 158 people, accounting for 37.62%, said that the school did not provide a convenient channel; 25 people said that they did not know whether the school provided a convenient channel, accounting for 5.95%. From the above data, the convenient channels provided by schools to students only account for 56.43%, indicating that schools still needed to work hard to provide convenient digital literature retrieval channels for students. Table 3 presents the digital literature retrieval promotion methods expected by students (multiple choice).

Table 3 shows that 310 people, accounting for 73.81%, hoped that Weibo and WeChat would be used as promotion methods; 267 people, accounting for 63.57%, hoped that e-books would be used as promotion methods; 183 people, accounting for 43.57%, wanted to use video as promotion method; 172 people, accounting for 40.95%, wanted to carry out novel promotion activities; 236 people, accounting for 56.19%, wanted to use film and television appreciation as the promotion method; 54 people, accounting for 12.86%, hoped that other methods can be used.

To sum up, schools should improve modern services and have achieved targeted promotion of digital literature retrieval by mastering students' digital literature retrieval habits. The future development environment would require more digital information. Colleges and universities must learn to find their own development direction in the new environment and continue to deepen and improve the promotion activities of digital document retrieval.

5. CONCLUSION

With the development of modern network technology and the multiplication of literature resources, coupled with the diversification of hardware equipment for collecting information, the production cost has been greatly reduced. Moreover, the trend of diversification of literature resources has become increasingly evident. Social networks on the Internet provide a more convenient and faster way to share literature resources, and numerous literature resources can be created by the public and widely disseminated. However, in most cases, these substantial literature resources only existed in their original form and have not been classified or managed effectively, thereby giving rise to the need for efficient management of literature resources. How to use these raw data resources to realise their

Table 3. Digital literature retrieval promotion methods expected by students (multiple choice)

Option	Number	Proportion
Weibo and WeChat	310	73.81%
e-book	267	63.57%
Video	183	43.57%
Novel activities	172	40.95%
Movies	236	56.19%
Other	54	12.86%

classification, identification and retrieval has become an important research content in the field of information processing. As a special type of data containing large amounts of textual information, the need to search datasets has led to the rapid development of various search algorithms over the past 40 years. Therefore, through the research on TRS IR, this study has proposed practical suggestions for the promotion of digital literature, which has important theoretical and practical significance. However, owing to the limitations of time and technology, this study did not analyse in detail the problems encountered in the research on the intelligent retrieval algorithm of digital literature promotion information and would further discuss this in the future.

CONFLICTS OF INTEREST

The authors of this publication declare there are no competing interests.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. Funding for this research was covered by the author(s) of the article.

REFERENCES

Ahmad, S., & Zhu, G. (2020). Politics and digital literature in the middle east: perspectives on online text and context. Palgrave Macmillan.

Cabanac, G., Frommholz, I., & Mayr, P. (2018). Bibliometric-enhanced information retrieval: Preface. *Scientometrics*, *116*(3), 1–3.

Chawla, S. (2017). Intelligent information retrieval using hybrid of fuzzy set and trust. *Oriental Journal of Computer Science and Technology*, *10*(2), 311–325. doi:10.13005/ojcst/10.02.09

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. doi:10.1016/j.joi.2017.10.003

Cunha, D., & Frota, W. N. (2017). On the path to a methodology for the critique of digital literature. *Literary and Linguistic Computing*, *32*(2), 225–233.

Djenouri, Y., Belhadi, A., & Belkebir, R. (2018). Bees swarm optimization guided by data mining techniques for document information retrieval. *Expert Systems with Applications*, *94*, 126–136. doi:10.1016/j.eswa.2017.10.042

Edmonds, C. (2019). Teaching literature with digital technology assignments, hetland tim. bedford/st. martin's, boston, ma (2017), 512 pp. *Computers and Composition*, *52*, 17–18. doi:10.1016/j.compcom.2019.02.005

Gianmaria, S., Georgeta, B., & Nicola, F. (2017). Semantic representation and enrichment of information retrieval experimental data. *International Journal on Digital Libraries*, *18*(2), 145–172. doi:10.1007/s00799-016-0172-8

Han, B., Chen, L., & Tian, X. (2018). Knowledge based collection selection for distributed information retrieval. *Information Processing & Management*, *54*(1), 116–128. doi:10.1016/j.ipm.2017.10.002

Hobbs, C., & Viinalass-Smith, S. (2018). Outside the margins and beyond the page: Complex digital literature, the new horizon for acquisition, conservation, curation and research. *Comma*, *2017*(1), 67–73. doi:10.3828/comma.2017.5

Holzapfel, A., Sturm, B. L., & Coeckelbergh, M. (2018). Information retrieval technology. *Transactions of the International Society for Music Information Retrieval*, *1*(1), 44–55. doi:10.5334/tismir.13

Kalyanasundaram, M., Abraham, S. B., & Ramachandran, D. (2017). Effectiveness of mind mapping technique in information retrieval among medical college students in Puducherry-A pilot study. *Indian Journal of Community Medicine: Official Publication of Indian Association of Preventive & Social Medicine*, *42*(1), 19–23. doi:10.4103/0970-0218.199793 PMID:28331249

Kamal, H. (2019). Blogging from egypt: Digital literature, 2005-2016. *Cairo Studies in English*, *2019*(2), 286–292. doi:10.21608/cse.2019.66678

Koerber, B. (2020). Blogging from egypt: Digital literature, 2005-2016, written by teresa pepe. *Journal of Arabic Literature*, *51*(1-2), 154–159. doi:10.1163/1570064x-12341403

Na, S. H., & Kim, K. (2017). Verbosity normalized pseudo-relevance feedback in information retrieval. *Information Processing & Management*, *54*(2), 219–239. doi:10.1016/j.ipm.2017.09.006

Onuoha, O. P. (2018). Diaspora digital literature: Role reversal and the construction of self in selected ikheloa's autobiographies. *International Journal of Pedagogy Innovation and New Technologies*, *5*(2), 33–42. doi:10.5604/01.3001.0012.8541

Rebecca, R. J., & Kalamani, S. (2020). Digital literature: A literary trend of the twenty first century. *International Journal of Advanced Research*, *8*(11), 725–728. doi:10.21474/IJAR01/12062

Taniguchi, Y. (2018). Content scheduling and adaptation for networked and context-aware digital signage: A Literature survey. *ITE Transactions on Media Technology and Applications*, *6*(1), 18–29. doi:10.3169/mta.6.18

Ubiparipovi, B., Matkovi, P., & Mari, M. (2020). Critical factors of digital transformation success: A literature review. *Ekonomika Preduzeca*, *68*(5-6), 400–415. doi:10.5937/EKOPRE2006400U

Varma, P. R., Sherly, E., & Mohan, K. (2017). Automated information retrieval model using fp growth based fuzzy particle swarm optimization. *International Journal of Information Technology and Computer Science*, *9*(1), 105–111. doi:10.5121/ijcsit.2017.9109