International Journal of Information Science and Management
Vol. 20, No. 2, 2022, 201-216
https://dorl.net/dor/20.1001.1.20088302.2022.20.2.12.7

Original Research

Discovering Research Topics from Medical Librarianship and Information Using Text Mining

Meisam Dastani
Ph.D. in knowledge and information science, Social Determinants of Health Research Center, Gonabad University of Medical Sciences, Gonabad, Iran.
meisam.dastani@gmail.com
ORCID iD: https://orcid.org/0000-0002-5631-539X

Afshin Mousavi Chelak
Associate Prof. Knowledge and Information Science Department, Payame Noor University, Tehran, Iran. mousaviaf@gmail.com
ORCID iD: https://orcid.org/0000-0002-5614-1334

Soraya Ziaei
Associate Prof., Knowledge and Information Science Department, Payame Noor University, Tehran, Iran.
Corresponding Author: soraya.ziaei@gmail.com
ORCID iD: https://orcid.org/0000-0003-3134-1468

Faeze Delghandi
Assistant Prof., Knowledge and Information Science Department, Payame Noor University, Tehran, Iran.
fdelghandi@gmail.com
ORCID iD: https://orcid.org/0000-0001-8540-7157

Received: 18 November 2020
Accepted: 20 July 2021

Abstract

An increasing number of articles published in different scientific fields makes it necessary to analyze the topics of these articles in specialized journals. For this purpose, topics published in the studies on medical librarianship and information in specialized journals were identified and analyzed in the present research. In the present study, an exploratory and descriptive approach was used to analyze medical librarianship and information articles published in specialized journals of this field from 1964 to 2019 by employing text-mining techniques. A latent Dirichlet Allocation (LDA) topic modeling algorithm was used to identify the published topics. Python programming language was also used to run text-mining algorithms. The findings of text mining and topic modeling showed that the following topics were published in medical librarianship and information: Patients' use of information resources (34%), Medical Librarianship and Information Services (18%), Scientometrics and bibliometrics (16.32%), Web-based treatment (15.47%), Information literacy and information skills (13.9%), and Trend and tweet analysis (1.92%). The publishing trend of articles in the medical librarianship and information indicates a change in research in the field.

Keywords*:* Medical Librarianship and Information, Content Analysis, Text Mining, Scientific Articles.

Introduction

Many scientific documents are produced every year by researchers, professors, and students of different universities, often containing valuable and useful materials and published in text formats. In addition, these scientific documents are published at different times and on different topics while, during the time, the publishing of topics changes. Therefore, studying macro and exploratory approaches is necessary for policymaking and strategic planning. Investigating the

process of publishing articles in scientific fields could present a suitable view of research and efforts in different fields of knowledge. By investigating the process of publishing articles in scientific fields, we could become aware of the subject changing over time. Because of the increasing number of scientific articles and the large numbers of published articles, it is impossible or very difficult to evaluate and review the details of articles and manually extract information and knowledge from this volume of texts. However, identifying patterns and extracting potential knowledge in large volumes of textual data is important in various scientific fields. Text mining is one of the analyzing methods of scientific texts and publishing (Hashimi, Hafez & Mathkour, 2015.

Text mining is also known as smart text analysis, text data mining, or knowledge discovery from a text generally refers to the knowledge extraction of desired and important information from unstructured text collection (Rodriguez-Esteban & Bundschus, 2016). Furthermore, identifying patterns and extracting potential knowledge in a large volume of text data is important in various scientific areas (Kao & Poteet, 2007). Salloum, Al-Emran, Monem & Shaalan (2018) described the objectives of text mining research as follows: 1) The use of text mining techniques to identify topics in scientific texts and the evolution of these topics, 2) The use of visualization tools to present each topic and the relationship between them as an appropriate method to help users determine relevant topics.

Medical librarianship and information is the subset of library and information science that is more than 100 years old. Medical Library Association (MLA) was established in 1898, proving this claim (Dimitroff, 1992). Medical information research is essential for medical librarians to develop new knowledge and evidence-based practice. Professional medical information specialists, who are members of the MLA, emphasize the importance and benefits of research in medical information (Lessick, Perryman, Billman, Alpi, De Groote & Babin Jr, 2016). Accordingly, specialists and researchers of medical librarianship and information conduct various researches and publish them in specialized journals in line with the objectives and missions of this scientific field. Many studies have employed text mining techniques to analyze scientific research and articles published in scientific journals and databases, including the publishing trend of articles in scientific journals and the most important topics related to insomnia and disorders (Lam, Lai, Wang, Lai, Hsu & Chung, 2012), identifying a comprehensive perspective on the evolution and development of information science (Yu, Pedrycz & Wang, 2017), analyzing the evolution of Mobile Health (Mhealth) research (Ozaydin et al., 2017), evaluating the reviews of articles presented at International AIDS Conferences over a period of more than 25 years to identify the trend of HIV terminology (Dancy-Scott et al., 2018), identifying the trend of research in the field of medical informatics to understand the current position of the medical informatics (Kim & Delen, 2018), recognizing the trend of published research in the field of anesthesia in reviews of articles presented in authoritative meetings (Rusanov, Miotto & Weng, 2018), analyzing the articles in the field of medical informatics (Saheb & Saheb, 2019), and clustering the articles of Information Science and Technology Research Center (SoleimaniNezhad, Salajegheh & Tayyebi Nia, 2019).

There are no studies on analyzing scientific texts in medical librarianship and information by employing text mining techniques; however, the studies that have evaluated and analyzed the articles published in this field have applied other methods such as bibliography and content analysis. The articles were published in Bulletin of the Medical Library Association between 1966 and 1990 (Dimitroff, 1992), the articles published in three journals named Medical

Information Service (MIS) in China, Igaku Toshokan (JJMLA) in Japan, and the Bulletin of the Medical Library Association (BMLA) in the USA between 1990 and 1992 (Zhang, 1994), the articles published in the Bulletin of the Medical Library Association and Journal of the Medical Library Association between 1991 and 2007 (Gore, Nordberg, Palmer & Piorun, 2009), the articles published in Medical Journal Reference Services Quarterly between 1982 and 2009 (Kenefick & Werner, 2011), the articles published in Bulletin of the Medical Library Association and Journal of the Medical Library Association between 1961 and 2010 (Funk, 2013) are such studies, which have often been conducted to identify topics, applied research methods, number of authors, number of citations, organizational dependence, and supporting organizations.

A review of the research background indicates that the analysis and evaluation of scientific publications are highly necessary for researchers, organizations, and scientific policymakers in the international arena. The rapid growth of scientific publications has created a large amount of scientific data. The solution to this challenge and encountering such a huge volume of publications is topic modeling and analyzing the keywords of articles using automated text mining. Text mining includes statistical methods investigating publications and documents to identify their themes (Abramson, Lees, Krzhizhanovskaya, Dongarra & Sloot, 2014).

According to this, it could be acknowledged that text mining is a reliable and practical tool for scientometric studies and could be a helpful tool for researchers to search in the scientific collection and identify subjects in the content of the text, the activities that traditional methods could not do. The result of text mining and topic modeling could analyze the relationship between the topic and its evolution over time and discover important topics and terms and trace them over time. In addition, it helps analyzers to understand better the communications and changes (Wang, Blei & Heckerman, 2012; Abramson et al., 2014; O'callaghan, Greene, Carthy & Cunningham, 2015). Also, its effectiveness in strategic decision-making and scientific policymaking at the macro level is very high (Cheng et al., 2018).

In this regard, the analysis and discovery of the structure of topics and identifying the topic evolution of published articles on medical librarianship and information over time are the main issues discussed in the present paper.

Materials and Methods

Sample and Data

This is applied research using text mining and an analytical approach. The statistical population is English articles published in medical librarianship and information in specialized journals between 1964 and 2019. By searching the list of journals of PubMed and Scopus databases, the topics of specialized journals of medical librarianship and information were extracted. Then, by searching the title of these journals in PubMed and Scopus databases on 20 January 2020, all published documents were extracted using the appropriate search strategy (Table 1).

PubMed is a database developed by the National Center for Biotechnology Information (NCBI) at the National Library of Medicine (NLM), one of the institutes of the National Institutes of Health (NIH). The database was designed to access citations (with abstracts) from biomedical journals. Also, the PubMed search engine (http://www.pubmed.gov), which provides access to the MEDLINE database operated by the United States National Library of

Medicine, has become the most popular online bibliographic resource for life scientists and clinicians (Vibert, Ros, Bigot, Ramond, Gatefin & Rouet, 2009; Canese, 2013). Because of the comprehensiveness of the Scopus database for different fields of sciences and the high indexing rate of articles, it is a suitable database for scientometrics studies (Mongeon & Paul-Hus, 2016).

Table1
The search strategy used in this research

Database	Search Strategies
PubMed	((((((((((("Journal of science librarianship"[Journal]) OR ("Journal of the European Association for Health Information and Libraries"[Journal])) OR "Biomedical digital libraries"[Journal]) OR "The journal of the Canadian Health Libraries Association"[Journal]) OR "Journal of the Medical Library Association: JMLA"[Journal]) OR ("Health information and libraries journal"[Journal])) OR "Journal of hospital librarianship"[Journal]) OR "Hospital libraries"[Journal]) OR "Health libraries review"[Journal]) OR "Bulletin of the Medical Library Association"[Journal]) OR ("Medical library and the historical journal"[Journal])) OR "Bulletin of the Association of Medical Librarians"[Journal]) OR "Medical libraries"[Journal]) OR "Journal of Medical Internet research"[Journal]
Scopus	(ISSN (1542-4073) OR ISSN (1540-9597))

Then the data extracted from the databases were integrated into a file and saved as CSV format containing 12819 records.

Subsequently, the items such as editor's notes, book reviews, content analysis, and letter to the editor were excluded because the purpose of the present investigation is the topic modeling of articles based on titles and abstracts. Also, it should be noted that these cases were not included in the survey articles; therefore, they did not have abstracts (Kim & Delen, 2018). After excluding these items, the sample data was reduced to 7599 articles.

Modeling and Data analysis process

The retrieved documents were then analyzed using text analysis. The text mining method used in this research is extracted from the designed framework by Zhang, Chen and Liu (2015), which was developed by Salloum et al. (2018). (Figure 1).

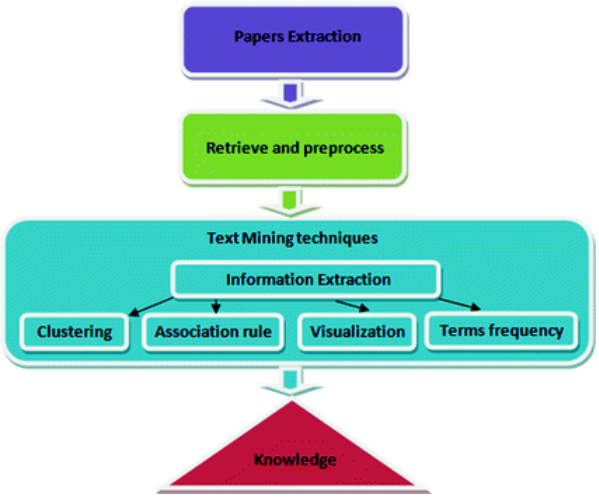

Figure 1: Conceptual model of text mining applied in this research

Then, the following steps were taken for text analysis: the pre-processing of texts, including document selection, extraction of words used in texts, and removing meaningless and stop words. In the next step, topic modeling and visualization algorithms were implemented using different text mining techniques, leading to extracting knowledge from texts and interpreting it (Figure 1).

In the present study, the Latent Dirichlet Allocation (LDA) algorithm was used for performing topic modeling operations (Blei et al., 2003). Topic modeling is a machine learning approach to discovering patterns or topics within a set of documents. LDA is one of the implementation methods in topic modeling (Blei, Ng & Jordan, 2003; Blei, 2012). LDA is one of the best algorithms extensively used and extremely effective in identifying related semantic topics in scientific texts (Griffiths & Steyvers, 2004) and has a better performance than many algorithms (Chang, Gerrish, Wang, Boyd-Graber & Blei, 2009). Therefore, LDA was selected for topic modeling in the present study. In addition to strengths, the LDA algorithm is limited in predicting the number of topics. The present study eliminated the number of predicted topics and the LDA limitation using the logarithmic (log) UMass Coherence criterion (Röder, Both & Hinneburg, 2015).

The elbow criterion was applied to identify the desired number of topics (Sbalchiero & Eder, 2020). By employing logarithmic (log) UMass Coherence and drawing the diagram of elbow criterion (Figure 2), from 6 to 40 topics could be selected for articles published in medical librarianship and information in specialized journals in this study; six topics were selected by investigating and interpreting topics with different values. It is noteworthy that selecting too many topics will lead to many small and similar topics (Greene, O'Callaghan . & Cunningham, 2014; Greene & Cross, 2017).

A larger number of topics also causes no additional topic information to be achieved. Moreover, it becomes more challenging to interpret topics because of the dispersion of keywords (Oliver, 2020). Then, the topics resulting from the implementation of the LDA algorithm were interpreted using the most important terms and articles of each topic. Python programming language and Gensim library have been applied to implement the topic modeling algorithm. The Gensim library is an open-source topic modeling tool with a simple syntax, which is compact, multi-purpose, and easy to develop and provides a variety of libraries for working with texts (Rehurek & Sojka, 2010). Figure 2 was obtained by implementing the logarithmic (log) UMass Coherence criterion based on the resulting data. The elbow point distinguishes the part with a steeper slope from the flatter part (Chen, Kou, Shang & Chen, 2015). In this diagram, *k* represents the number of topic categories.

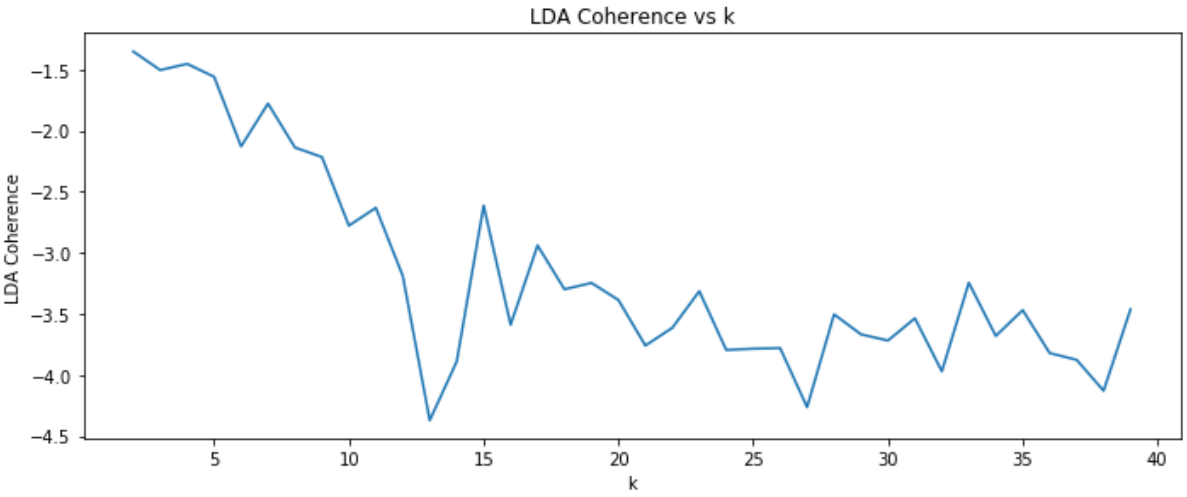

Figure 2: The elbow diagram obtained from the UMass Coherence algorithm's implementation to select the number of topics

The LDA algorithm applied in the present research determines the optimal number of topics, the frequency distribution of each document in the selected topics, and the list of keywords related to each topic. However, it does not have the capability of automatic labeling. Hence, topic labels are defined and specified manually (Blei, 2012).

Results

Publishing Trend of Articles in the field of Medical Librarianship and Information

Figure 3 presents the publishing trend of articles in the field of medical librarianship and information over time. The trend of publishing articles grew in the last ten years, and the highest number of articles were published in 2019.

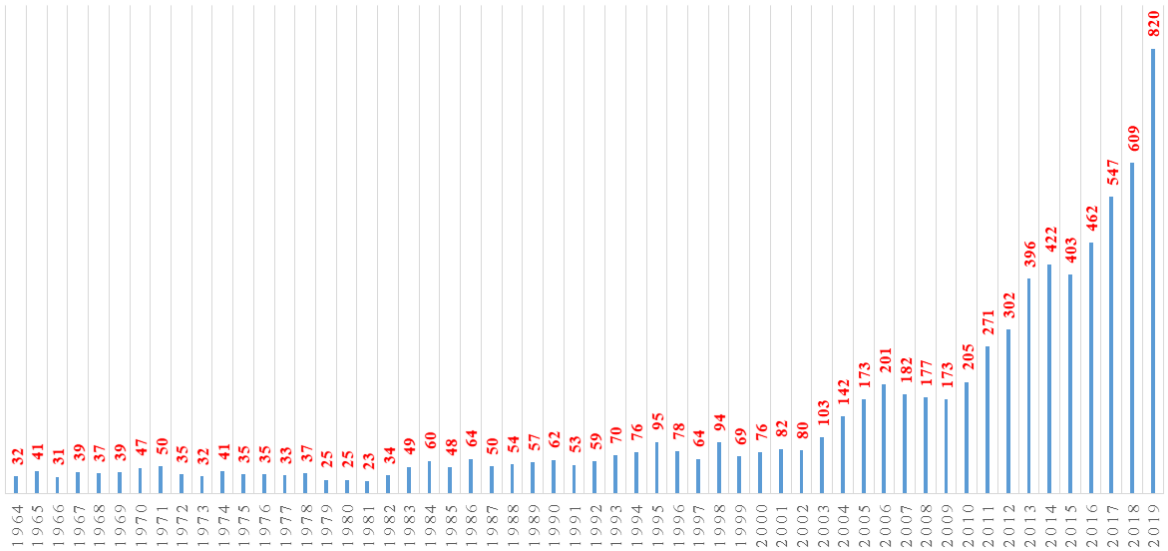

Figure 3: The publishing trend of articles in the field of medical librarianship and information over time

The topic of Articles Published in the field of Medical Librarianship and Information

The results obtained from implementing the Latent Dirichlet Allocation (LDA) topic-modeling algorithm are shown in Figure 4. It also illustrates 10 important words of each topic in the form of a word cloud with a specific color. The figure shows the trend of terms used in this research. Word clouds provide a unique way to summarize the content of text documents. In a word cloud, the size of a word in the cloud is proportional to its importance and application in the whole text collection (Cui, Wu, Liu, Wei, Zhou & Qu, 2015). For instance, in topic number one, the words Tweet and Post were the most important words used in the texts of articles related to this topic, and the words health and patient were the most important words used in the texts of articles related to the topic number two.

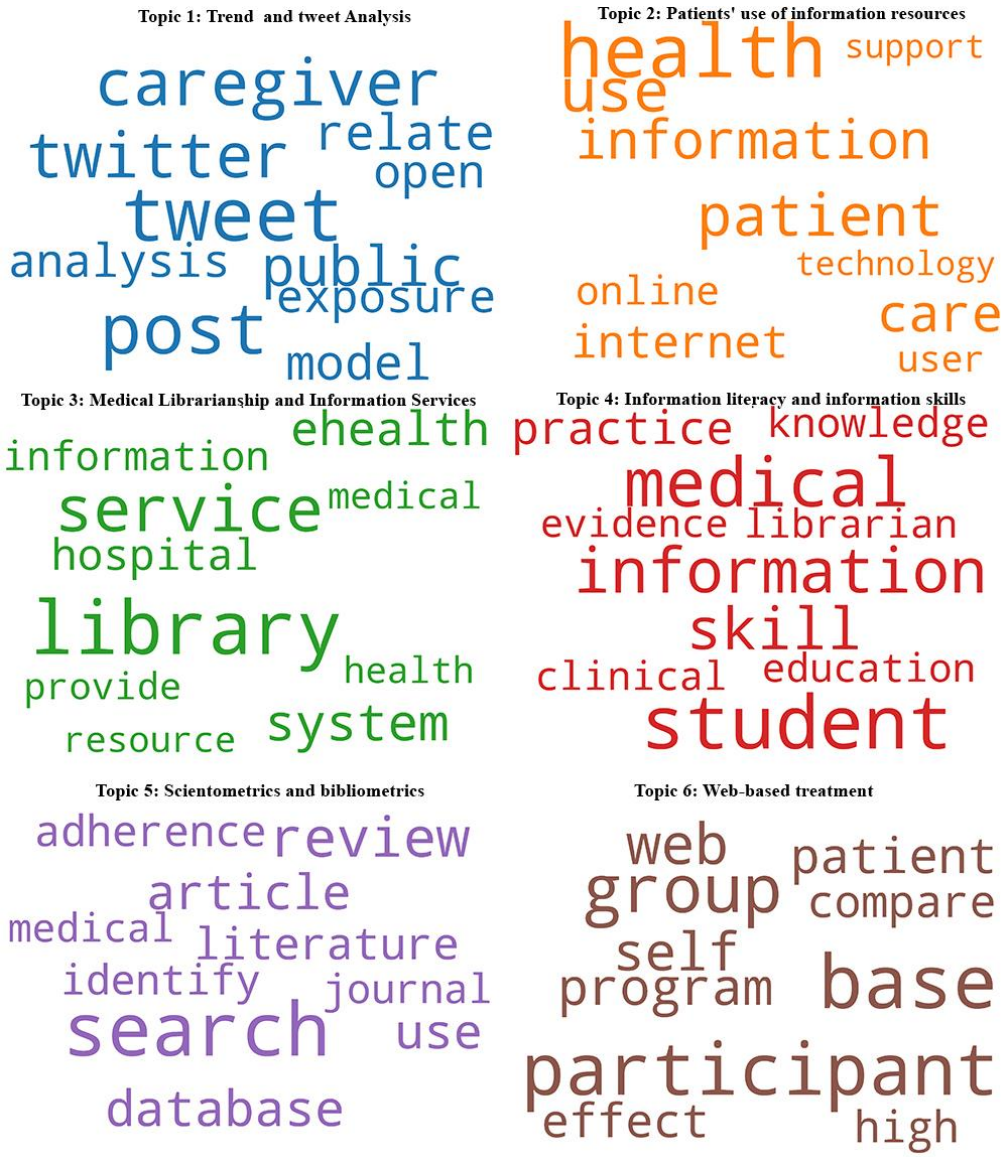

Figure 4: Word cloud of topics obtained from topic modeling algorithm of articles in the field of medical librarianship and information

Figure 5 shows the publishing rate of articles on each topic. Topic 2, which deals with "Patients' Use of Information Resources", had the highest publishing rate of 34%, while "Trend

and Tweet Analysis" had the lowest publishing rate of 1.92%.

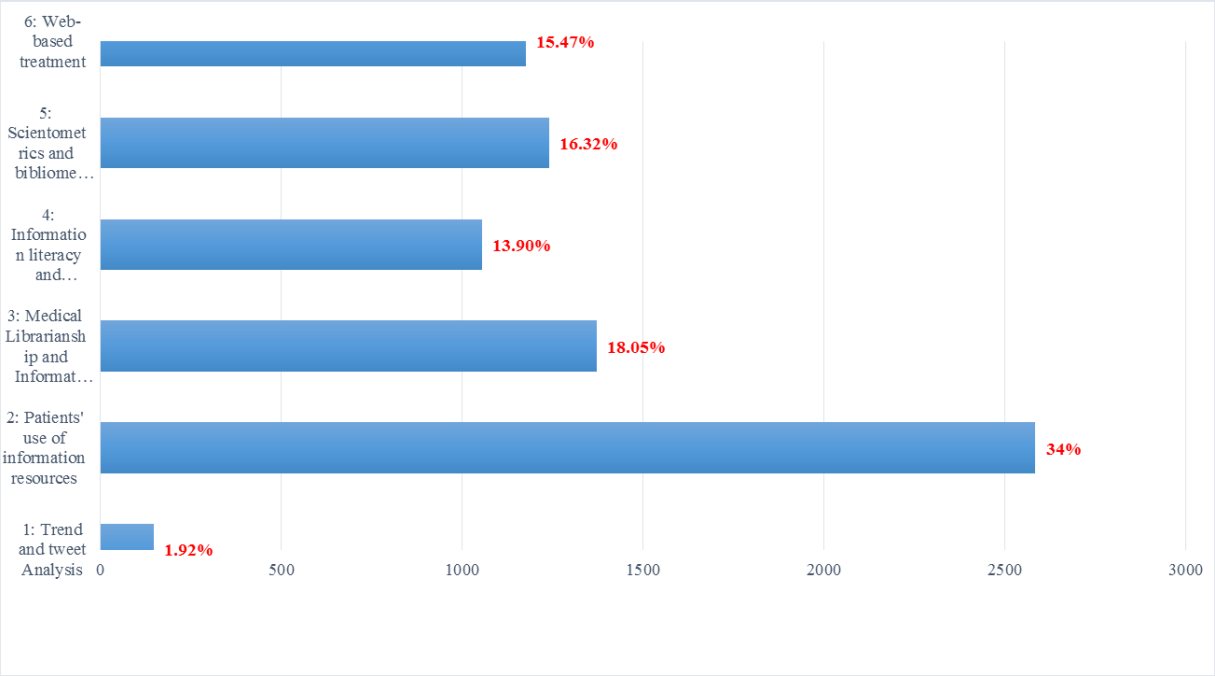

Figure 5. The percentage of articles published on each of the topics of the medical librarianship and information

Publishing Trend of Topics of Articles in the Field of Medical librarianship and information

Figure 6 indicates the publishing trend of different topics of articles in medical librarianship and information. As can be seen, the peaking of publication of different topics in this field started in 2010, and the highest publication of articles on this topic was in 2019. The data in Figure 6 indicate that the highest publication rate is related to "Patients' use of information resources". After this topic, "Medical Librarianship and information services" had the highest publication rate, which was slow and the same publishing trend over time. "Web-based treatment" is also a topic with an increasing trend since 2010. The highest rate of publication on this topic was in 2010.

The other topics shown in Figure 6 had a relatively slow and relatively the same publishing trend over time.

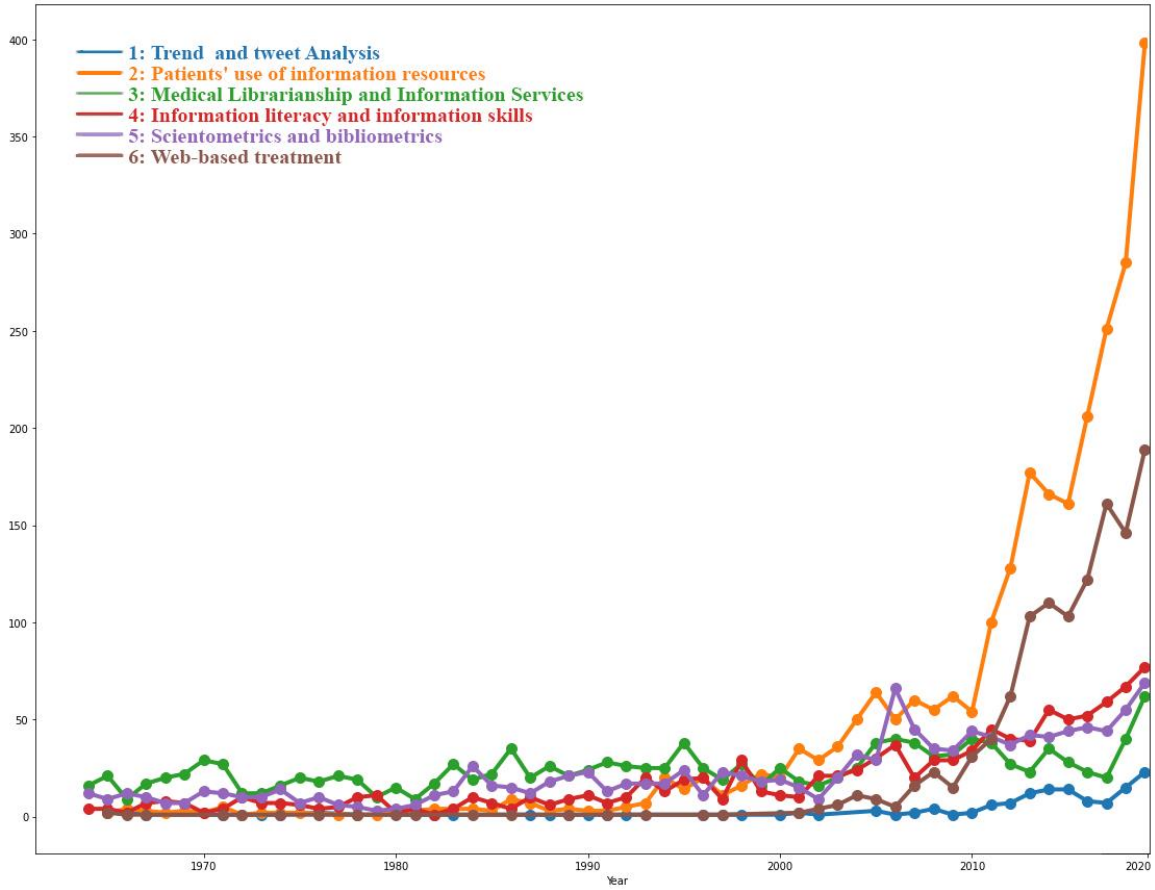

Figure 6: Publishing trend of topics of articles in the field of medical librarianship and information

Discussion

Topic modeling acts as a text mining tool for processing, management, organizing, and knowledge extraction. It is commonly used to identify basic "topics" in texts (Lamba & Madhusudhan, 2019) and can present a good view of a large collection in terms of the collection as a whole, independent documentation, and relationships between documents (Jelodar et al., 2019). The present study demonstrates an obvious illustration of the situation of articles published in the field of medical librarianship and information from 1964 to the end of 2019 in specialized journals of medical librarianship and information. The results of the topic modeling of the research identified six topics of patients' use of information resources, medical libraries and information services, scientometrics and bibliometrics, web-based treatment, information literacy and information skills, trend, and tweet analysis in order of highest number of publications for articles published in journals.

The topics obtained in the present study are subsets of main subjects in library and information science - library science, information retrieval, and bibliographies (Larivière, Sugimoto & Cronin, 2012).

Also, the obtained subjects of the present study are aligned to the three values of the library and information profession - 1) Access to information for all those who need to use, 2) Improving literacy, and 3) Preservation of accumulated past wisdom (Groen, 2007). The topics obtained in the present study are all in line with these values, but other topics were obtained due to the specialized and medical nature of the library and information. As the results show, the highest publishing rate is related to the topic of “Patients' use of information resources",

which shows the importance and high value of this subject among researchers in medical librarianship and information. This topic has been growing since 2010, and the peak of publishing was in 2019.

Kurata et al. (2018) have also discussed searching for information on the web, search behavior, and healthcare information as topics of interest to librarians in recent years. Figuerola, García, Marco and Pinto (2017) considered the beginning and growth of this field to be affected by the advent of the Internet. Medical librarians are also more concerned with digital information than physical resources and packages in the past and are looking for evidence-based physical resources and packages (Funk, 2013). Moreover, access to high-quality information in health promotion is of great importance, and accordingly, the role of medical librarians is changing. Professionals of this area can be useful for decision-making in line with health improvement due to their skill in accessing and using the correct information. Medical librarians can also effectively educate users to access appropriate information (Homan & McGowan, 2002).

In this regard, Gavgani and Mohan (2008) stated that one of the tasks of medical librarians is to provide health services and support the information needs of consumers. Library science is naturally a field of socialization that always adapts to its user community's changing needs and behaviors. Medical Librarians also know to organize, search, find, locate, and deliver accurate, reliable, and related information. They know how and when to provide information to physicians and patients and find useful information from this large volume of information.“Medical Librarianship and information services” is another topic with the highest publishing rate, with a slow publishing trend over time. The results of a study conducted by Dimitroff (1992) showed that the articles published in the Bulletin of the Medical Library Association between the years 1966 and 1990 were mostly related to applied topics of library and information such as management, providing public and technical services, materials and collections and behavior, and users' attitudes and opinions.

Also, using the continually changing technology, managing the libraries has changed (Funk, 2013). Gore et al. (2009) showed that the interest and emphasis of medical librarianship and information researchers have shifted towards library users and public services; however, they had focused on the physical library (i.e., used systems, daily operations) in the past years. As can be seen, studies of library materials and collections continue to account for a significant percentage of published research. Moreover, studies that evaluate the way of employing library resources are increasingly common. Such a change indicates more customer access via the Internet. The present study is on the medical librarianship and information and shows the efforts of the researchers to answer the questions related to libraries and the services they provide in such an environment. Its publishing trend also shows that this is one of the most important issues over time, from the past to the present.

“Scientometrics and bibliometrics” was also the topic with a slow publishing trend over time, and the highest number of articles in this field was published in 2008. This topic deals with the quantitative study of information and analyzing scientific articles, publications, and journals. The majority of articles on this topic are in scientometrics and bibliometrics. Figuerola et al. (2017) showed that the pick of publishing articles on this topic was on the LISA database between 2003 and 2013. The studies conducted by other researchers have also indicated that scientometrics and bibliometrics were in the high interest of librarianship researchers in recent years (Lamba & Madhusudhan, 2019; Kurata et al., 2018; Olmeda-Gómez, Ovalle-Perandones

& Perianes-Rodríguez, 2017). Zhang (1994) showed that information retrieval and evaluation of books and journals were the most important topics published in medical librarianship between 1990 and 1992. Kenefick and Werner (2011) also showed that dissemination or information retrieval, reference, bibliographic instruction, administration and management, and bibliographies or literature reviews were the most common topics published in the Medical Reference Services Quarterly journal between 1982 and 2009.

"Information literacy and information skills" is one of the topics that researchers have considered in the field of medical librarianship and information in recent years. Figuerola et al. (2017) reported the beginning of the publication of articles on this topic in 1993 after the advent of the Internet, and they also showed that its growth was since the year 2002. This topic has also experienced a higher growth rate due to the development of online learning platforms and new technologies. The development of the web, social networks, and media communications have also been other reasons for this growth (Hobbs & Jensen, 2009). Lamba and Madhusudhan (2019) conducted topic modeling on the DESIDOC Journal of Library and Information Technology in India and concluded that information literacy is among the topics that have been extensively published in the journal over the past decade. Kurata et al. (2018) indicated that most articles on the users' use of computers have focused on the search process. Moreover, searching for information on the web, search behavior, and information literacy were among the topics that have received considerable attention recently.

"Web-based treatment" is also an important topic. As the results show, the growth of this topic has started since the year 2010, and the pick of publishing articles on this topic was in the year 2019, which indicates that researchers have considered it in recent years. This high-interest rate could be due to the growth and development of Internet-based services. It can be said that this topic is an interdisciplinary issue in line with the field of health information technology and medical informatics. In some specialties, the medical librarianship and information experts act similarly to their colleagues in other fields. Therefore, the library's role in using information technology in accessing, owning, and publishing the information has brought it closer to the field of medical informatics (Frisse, Braude, Florance & Fuller, 1995). Kim and Delen (2018) showed that the topics related to Internet methods and the knowledge presentation were more important for researchers in medical informatics, and there were many themes around these two topics for treatment and better diagnosis of diseases that used the Internet.

"Trend and Tweet Analysis" has the lowest number of publications, and articles related to this topic have been published since 2010. The published topics in all fields of science are evolving, new topics are emerging, and old topics are becoming abolished over a period. Topic modeling not only assists the researchers in identifying topics related to each of the scientific fields but also helps them recognize new and growing topics.

The present research has obtained the topics of articles in the field of medical librarianship and information in specialized journals using text-mining techniques and topic modeling tools and has identified the publishing rate of the articles in each of these topics. In addition, the publishing trend and growth of these topics were determined from the past to the present. The topics of patients' use of information resources and web-based treatment have grown increasingly in recent years, and the highest number of articles on these topics were published in 2019. The publishing trend of these topics shows a change in the medical librarianship and information studies.

Conclusion

According to the results of the present study, text-mining techniques in the identification of the published topics in the field of medical library and information can help researchers select novel and up-to-date topics. In addition, they are useful for managers and policy-makers of related journals in determining research priorities so that they pay more attention to some topics that have been less used in related articles.

Furthermore, according to the findings of the present investigation, it is suggested that the managers of journals, the focuses of which are on publishing articles in different subject areas, use automatic text mining techniques to identify the published articles as well as their publishing rate in each of these topics, and then, based on the obtained results, take action for planning and policy-making to publish articles in various subject areas in the journal

Limitations and Future Works

One of the limitations of the present research was the studied population. The studied articles were published in the specialized journals of medical librarianship and information in PubMed and Scopus databases in the present research. A small number of articles on medical librarianship and information may be published in non-specialized journals, and since the review of all of these journals was not manually possible, these articles were not included in the study.

The methodology of studies on topic modeling was another limitation of the present research, including identifying an appropriate number of topics for the articles before performing the Latent Dirichlet Allocation. The incompetence of the Dirichlet topic distribution is the correlation among topics, manual interpretation, and labeling of topics. Although some topics were reasonably straightforward to label, others proved more difficult to ascertain the content or methodological relationship that connected the words and the representative articles.

According to the results obtained in this study for future works, authorities and policymakers in the medical librarianship and information could determine research priorities of the scientific field, and researchers could use the research methodology to precisely and accurately analyze the published articles of each of the topics in the present study. The methodology of this research can also be employed to discover published topics in other scientific fields.

Acknowledgment

The data used in this article were obtained from a Ph.D. dissertation at the Payame Noor University of Mashhad. We thank all those who have helped in this research.

References

Abramson, D., Lees, M., Krzhizhanovskaya, V. V. Dongarra, J. J., & Sloot, P. M. (2014, January). Big data meets computational science, preface for ICCS 2014. In *ICCS* (pp. 1-7).

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84. https://doi.org/10.1145/2133806.2133826

Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3, 993-1022. Retrieved from https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf

Canese, K. & Weis, S. (2013). PubMed: The bibliographic database. The NCBI Handbook, 2(1). Retrieved from http://www.ehu.eus/biofisica/juanma/mbb/pdf/pubmed_intro.pdf

Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J. L. & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. Paper presented at *the Advances in neural information processing systems, 2*. Retrieved from http://users.umiacs.umd.edu/~jbg/docs/nips2009-rtl.pdf

Chen, K., Kou, G., Shang, J. & Chen, Y. (2015). Visualizing market structure through online product reviews: Integrate topic modeling, TOPSIS, and multi-dimensional scaling approaches. *Electronic Commerce Research and Applications*, 14(1), 58-74. https://doi.org/10.1016/j.elerap.2014.11.004

Cheng, X., Shuai, C., Liu, J., Wang, J., Liu, Y., Li, W. & Shuai, J.. (2018). Topic modeling of ecology, environment and poverty nexus: An integrated framework. *Agriculture, Ecosystems & Environment*, 267, 1-14. https://doi.org/10.1016/j.agee.2018.07.022

Cui, W., Wu, Y., Liu, S., Wei, F., Zhou, M. X., & Qu, H. (2010). Context preserving dynamic word cloud visualization. Paper presented at the 2010 IEEE Pacific Visualization Symposium (PacificVis).

Dancy-Scott, N., Dutcher, G. A., Keselman, A., Hochstein, C., Copty, C., Ben-Senia, D., ... & Choi, J. J. (2018). Trends in HIV terminology: text mining and data visualization assessment of International AIDS conference abstracts over 25 years. *JMIR Public Health And Surveillance*, 4(2), e50. https://doi.org/10.2196/publichealth.8552

Dimitroff, A. (1992). Research in health sciences library and information science: a quantitative analysis. *Bulletin of the Medical Library Association*, 80(4), 340.

Figuerola, C. G., García Marco, F. J. & Pinto, M. (2017). Mapping the evolution of library and information science (1978–2014) using topic modeling on LISA. *Scientometrics*, 112(3), 1507-1535. https://doi.org/10.1007/s11192-017-2432-9

Frisse, M. E., Braude, R. M., Florance, V. & Fuller, S. (1995). Informatics and medical libraries: changing needs and changing roles. *Academic medicine*: *journal of the Association of American Medical Colleges*, 70(1), 30-35. https://doi.org/10.1097/00001888-199501000-00010

Funk, M. E. (2013). Our words, our story: A textual analysis of articles published in the Bulletin of the Medical Library Association/Journal of the Medical Library Association from 1961 to 2010. *Journal of the Medical Library Association: JMLA*, 101(1), 12-20. https://dx.doi.org/10.3163%2F1536-5050.101.1.003

Gavgani, V. Z. & Mohan, V. V. (2008). Application of Web 2.0 Tools in Medical Librarianship to Support Medicine. *Webology*, 5(1). Retrieved from https://www.webology.org/data-cms/articles/20200515034535pma53.pdf

Gore, S. A., Nordberg, J. M., Palmer, L. A. & Piorun, M. E. (2009). Trends in health sciences library and information science research: an analysis of research publications in the Bulletin of the Medical Library Association and Journal of the Medical Library Association from 1991 to 2007. *Journal of the Medical Library Association: JMLA*, 97(3), 203-211. https://doi.org/10.3163/1536-5050.97.3.009

Greene, D. & Cross, J. P. (2017). Exploring the political agenda of the European parliament using a dynamic topic modeling approach. *Political Analysis*, 25(1), 77-94. https://doi.org/10.1017/pan.2016.7

Greene, D., O'Callaghan, D., & Cunningham, P. (2014, September). How many topics? stability analysis for topic models. In *Joint European conference on machine learning and knowledge discovery in databases* (pp. 498-513). Springer, Berlin, Heidelberg.

Griffiths, T. L. & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101(suppl 1), 5228-5235. https://doi.org/10.1073/pnas.0307752101

Groen, F. K. (2007). *Access to medical knowledge: Libraries, digitization, and the public good*. Scarecrow Press.

Hashimi, H., Hafez, A. & Mathkour, H. (2015). Selection criteria for text mining approaches. *Computers in Human Behavior*, 51(part B), 729-733. https://doi.org/10.1016/j.chb.2014.10.062

Hobbs, R. & Jensen, A. (2009). The past, present, and future of media literacy education. *Journal of Media Literacy Education*, 1(1). Retrieved from https://digitalcommons.uri.edu/jmle/vol1/iss1/1

Homan, J. M. & McGowan, J. J. (2002). The Medical Library Association: promoting new roles for health information professionals. *Journal of the Medical Library Association: JMLA*, 90(1), 80-85.

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y. & Zhao, L. (2019). Latent Dirichlet Allocation (LDA) and Topic modeling: models, applications, a survey. *Multimedia Tools Applications*, 78(11), 15169-15211. https://doi.org/10.1007/s11042-018-6894-4

Kao, A. & Poteet, S. R. (Eds.). (2007). *Natural language processing and text mining*. Springer Science & Business Media.

Kenefick, C. & Werner, S. E. (2011). Bibliometric study of medical reference services quarterly, 1982–2009. *Medical Reference Services Quarterly*, 30(1), 1-11. https://doi.org/10.1080/02763869.2011.540204

Kim, Y. M. & Delen, D. (2018). Medical informatics research trend analysis: A text mining approach. *Health informatics journal*, 24(4), 432-452. https://doi.org/10.1177%2F1460458216678443

Kurata, K., Miyata, Y., Ishita, E., Yamamoto, M., Yang, F., Iwase, A. & Technology. (2018). Analyzing library and information science full-text articles using a topic modeling approach. *Proceedings of the Association for Information Science*, 55(1), 847-848. https://doi.org/10.1002/pra2.2018.14505501143

Lam, C., Lai, F. C., Wang, C. H., Lai, M. H., Hsu, N. & Chung, M. H. (2016). Text mining of journal articles for sleep disorder terminologies. *PloS ONE*, 11(5), e0156031. https://doi.org/10.1371/journal.pone.0156031

Lamba, M. & Madhusudhan, M. (2019). Mapping of topics in DESIDOC Journal of Library and Information Technology, India: A study. *Scientometrics*, 120(2), 477-505. https://doi.org/10.1007/s11192-019-03137-5

Larivière, V., Sugimoto, C. R. & Cronin, B. (2012). A bibliometric chronicling of library and information science's first hundred years. *Journal of the American Society for Information Science and Technology*, 63(5), 997-1016. https://doi.org/10.1002/asi.22645

Lessick, S., Perryman, C., Billman, B. L., Alpi, K. M., De Groote, S. L. & Babin Jr, T. D. (2016). Research engagement of health sciences librarians: a survey of research-related activities and attitudes. *Journal of the Medical Library Association: JMLA*, 104(2), 166-173. https://doi.org/10.3163/1536-5050.104.2.015

Mongeon, P. & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213-228. https://doi.org/10.1007/s11192-015-1765-5

O'callaghan, D., Greene, D., Carthy, J. & Cunningham, P. (2015). An analysis of the coherence of descriptors in topic modeling. *Expert Systems with Applications*, 42(13), 5645-5657. https://doi.org/10.1016/j.eswa.2015.02.055

Oliver, W. & Schubert, D. (2020). The Symbolic Power of the Research Excellence Framework. Evidence from a case study on the individual and collective adaptation of British Sociologists. Retrieved from yun.ir/sn0mi6

Olmeda-Gómez, C., Ovalle-Perandones, M.-A. & Perianes-Rodríguez, A. (2017). Co-word analysis and thematic landscapes in Spanish information science literature, 1985–2014. *Scientometrics*, 113(1), 195-217. https://doi.org/10.1007/s11192-017-2486-8

Ozaydin, B., Zengul, F., Oner, N. & Delen, D. (2017). Text-mining analysis of mHealth research. *mHealth*, 3, 53. https://doi.org/10.21037/mhealth.2017.12.02

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.*

Röder, M., Both, A. & Hinneburg, A. (2015, February). Exploring the space of topic coherence measures. In *Proceedings of the eighth ACM international conference on Web search and data mining* (pp. 399-408), Shanghai, China.

Rodriguez-Esteban, R. & Bundschus, M. (2016). Text mining patents for biomedical knowledge. *Drug Discovery Today*, 21(6), 997-1002. https://doi.org/10.1016/j.drudis.2016.05.002

Rusanov, A., Miotto, R. & Weng, C. (2018). Trends in anesthesiology research: A machine learning approach to theme discovery and summarization. JAMIA Open, 1(2), 283-293. https://doi.org/10.1093/jamiaopen/ooy009

Saheb, T. & Saheb, M. (2019). Analyzing and visualizing knowledge structures of health informatics from 1974 to 2018: A bibliometric and social network analysis. *Healthcare Informatics Research*, 25(2), 61-72. https://doi.org/10.4258/hir.2019.25.2.61

Salloum, S. A., Al-Emran, M., Monem, A. A. & Shaalan, K. (2018). *Using text mining techniques for extracting information from research articles. In Intelligent natural language processing: Trends and applications* (pp. 373-397). Springer, Cham.

Sbalchiero, S. & Eder, M. (2020). Topic modeling, long texts and the best number of topics. Some Problems and solutions. *Quality & Quantity*, 54(4), 1095-1108. https://doi.org/10.1007/s11135-020-00976-w

SoleimaniNezhad, A., Salajegheh, M. & Tayyebi Nia, E. (2019). Clustering scientific articles based on the k_means algorithmcase study: Iranian research institute for information science and technology (IranDoc). *Iranian Journal of Information Processing and Management.* 34(2), 871-896. [in Persian]

Vibert, N., Ros, C., Bigot, L. L., Ramond, M., Gatefin, J. & Rouet, J. F. (2009). Effects of domain knowledge on reference search with the PubMed database: An experimental study. *Journal of the American Society for Information Science and Technology*, 60(7), 1423-1447. https://doi.org/10.1002/asi.21078

Wang, C., Blei, D. & Heckerman, D. (2012). Continuous time dynamic topic models. *arXiv preprint arXiv:1206.3298.*

Yu, D., Xu, Z., Pedrycz, W. & Wang, W. (2017). Information Sciences 1968–2016: A retrospective analysis with text mining and bibliometric. *Information Sciences*, 418, 619-634. https://doi.org/10.1016/j.ins.2017.08.031

Zhang, H. (1994). A bibliometric study of the articles on medical librarianship: Comparison of medical library periodicals published in three countries. *Igaku Toshokan*, 41(1), 84-97. https://doi.org/10.7142/igakutoshokan.41.84

Zhang, Y., Chen, M. & Liu, L. (2015, September). A review on text mining. In *2015 6th IEEE International Conference on Software Engineering and Service Science (ICSESS)* (pp. 681-685). IEEE.