Network analysis and scientific mapping of the e-learning literature from 1995 to 2018

Mohammad Bagher Negahban  
Shahid Bahonar University of Kerman, Iran  
Shiraz University of Medical Science, Shiraz, Iran  

Nahid Zarifsanaiey  
Shiraz University of Medical Sciences, Shiraz, Iran  
Virtual University of Medical Sciences, Tehran, Iran

Knowledge Management & E-Learning: An International Journal (KM&EL)  
ISSN 2073-7904

Recommended citation:  
https://doi.org/10.34105/j.kmel.2020.12.014
Network analysis and scientific mapping of the e-learning literature from 1995 to 2018

Mohammad Bagher Negahban*  
Department of Information Science & Knowledge Studies  
Shahid Bahonar University of Kerman, Iran  
Shiraz University of Medical Science, Shiraz, Iran  
E-mail: bm.negahban@gmail.com

Nahid Zarifsanaiey  
Department of E-learning, Virtual School  
Comprehensive Center of Excellence for e-Learning in Medical Sciences  
Shiraz University of Medical Sciences, Shiraz, Iran  
Department of Medical Education  
Virtual University of Medical Sciences, Tehran, Iran  
E-mail: nzarifsanaee@gmail.com

*Corresponding author

Abstract: This study aims to explore the research trends in the field of e-learning by analyzing the relevant publications using network analysis and scientific mapping. The data were extracted from the Web of Science database from 1995 to 2018, revealing a growing number of articles published in e-learning. The keywords “e-learning” and “Internet” ranked first and second as the most frequently used terminology. “Internet”, “Blended Learning” and “E-Learning” are recognized as three main categories in the field. The findings also reveal that the research related to Internet, distance learning, learning management system, MOOC and mobile learning has had the greatest impact in the field of e-learning. Further, some topics (e-learning, education, MOOC, learning, online learning, blended learning, learning management system) serve as bridge to connect different topics. The other highly used terms, such as “information technology”, “healthcare”, “management” and “blended learning” represent the emerging topics in this field. In particular, e-learning in health or medical care has been an actively promoted topic in this field.

Keywords: e-Learning; Citation analysis; Network analysis; Scientific map

Biographical notes: Dr. Mohammad Bagher Negahban Assistant Professor in Information Science at the Shahid Bahonar University Of kerman and He is also an Academic Fellow of Shiraz University of Medical Science, Iran. Her research interests focus on the Information Science, Scientometrics and e-Learning.

Dr. Nahid Zarifsanaiey Associate Professor in e-Learning at the Virtual School, Comprehensive Center of Excellence for e-Learning in Medical Sciences at Shiraz University of Medical Sciences, Shiraz, Iran and she is also an Academic Fellow at Department of Medical Education, Virtual University of Medical Sciences, Iran.
1. Introduction

Network analysis and scientific mapping of different disciplines and tracking their latest developments is a topic of interest to most researchers and scholars. Nowadays scientific output, as an indicator of activities in a country’s scientific system, has become the focus of attention for policy makers at national and international levels. Although scientific production alone cannot be regarded as a definite marker of comprehensive scientific growth and development, many R&D decisions today are based on mapping and measuring the countries’ scientific output. In view of its dynamic nature and its direct link to technology and scientific development, electronic learning has emerged as an important area of science, which is offered as an academic discipline in universities up to the doctoral level. Today, the development of electronic teaching/learning and virtual space play a critical role in the development of scientific knowledge, and the research activities in this field have progressed in parallel to the activities in other disciplines.

The researchers in this field have published a substantial number of scholarly papers in reliable scientific databases, which have led to the formation of its scientific structure. Mapping the scientific and social structure of e-learning presents a valuable source of information about the position of each of its subfields, signifying its overall power. Scientific mapping is valuable and can be subjected to scientometric data analysis. Accordingly, evaluating the scientific output is a factor that facilitate the assessment of the achieved efficiency in this particular discipline. Therefore, the present research analyzes the studies conducted in the field of e-learning by means of network analysis method.

Network analysis refers to the study of the characteristics of networks and the relationship between their vertices and segments based on the network theory or graph theory. Various measures are used to analyze networks, the most important of which are closeness and betweenness centrality measures. The centrality index represents the social power of a vertex and its ranking in the network based on its role and position in the network communications. Closeness centrality measure is calculated based on the Geodesic distance, and measures the distance between one node to the next. It represents the average length of the shortest paths between each individual and other individual on the network. In other words, closeness centrality is a point that is, on average, close to all other points. The closer a point is to the center, the stronger it is. Betweenness centrality is measured by quantifying the number of times a node that appears along the shortest path between two other nodes. This index measures centrality by examining the extent to which a particular actor is placed among other diverse actors in the network. The betweenness centrality measure identifies the position of an entity on a network according to its ability to communicate with other pairs or to serve as a bridge from one node to another on the same network.

This kind of network analysis can help to find the unknown connections in an area of science, spreading an idea over a specific time period, revealing the trends in a particular field, identifying the important and prominent topics in a field, and detecting the areas of interest and dominant concepts in researchers’ work. In this respect, a relevant concept is co-occurrence analysis, which is a measure that focuses on the most frequently used words to identify the most important research topics in any scientific field. In other words, the frequency of a word is considered as an indicator of its importance, the attention or emphasis that is placed on that word or idea (Leydesdorff,
Co-occurrence analysis of keywords is usually defined as an assessment of index terms, which are assigned to a document by the indexing services or by the document creators.

These network analysis methods are critical in identifying the emerging trends and their dynamics in the rapidly evolving fields like electronic learning. In recent decades electronic learning has received considerable attention from scientific and academic circles. It focuses on utilizing different components of information and communication technology, especially the Internet and the Web, in organizing and managing educational systems. The term e-learning refers to any form of learning and teaching that takes place through a variety of electronic media. In fact, it involves the use of technological innovations to enable and empower individuals to learn and teach regardless of time and location. In this respect, such terms as virtual learning, machine learning, Internet and Web learning, online learning, blended learning and distance learning are among the most commonly used terms, which will be analyzed in this study.

The main objectives of this research are to determine the growth trend of scientific production; to identify emerging and prominent subject areas; and to identify and analyze the co-occurrence network of the e-learning based on the closeness and betweenness centrality measures.

This study aims to answer the following four research questions (RQs):

**RQ1:** How is the growth trend of scientific production in the field of E-learning?

**RQ2:** What are the emerging and prominent subject areas in the field of e-learning?

**RQ3:** What topics have high impact in the e-learning field?

**RQ4:** What topics serve as a bridge from one topic to another in the e-learning field?

2. Literature review

Numerous researchers have conducted scientific mapping and network analysis in a wide range of disciplines. For instance, (Lee & Su, 2011) analyzed 223 highly cited articles from 142 institutions in 269 countries, in the area of electrical conducting polymer nanocomposites, with the aim of presenting a quantitative map of scientific research. In an article entitled “Mapping knowledge domains of Chinese digital library research output, 1994-2010”, Zhao and Zhang (2011) attempted to find the intellectual paradigms in the field of digital libraries. (Borrett et al., 2013) published an article entitled “The rise of network ecology: Maps of the topic diversity and scientific collaboration”. (Xie, 2014) presented an article entitled “Hotspots of Ecological and Environmental Risk Research in China Based on Co-word Analysis”, which examined the hotspots in the area of environmental and ecological risk research in China, using co-word analysis of keywords. In a research study, (Surjandari et al., 2015) conducted a co-word analysis of the University of Indonesia’s scientific output between 2010 and 2015. In a study entitled “Research patterns and trends of Recommendation System in China, using co-word analysis”, (Zhang et al., 2015) examined these patterns and trends in China during 2003-2014. In a research study entitled “Mapping the World of Science”, (White & Garfield, 1998) performed this procedure, using SCI. In another research entitled “Global Map of science Based on the ISI subject categories”, (Leydesdorff et al., 2013) initially analyzed the differences and similarities amongst the thematic categories of Web of Science in versions 4 and 5. (Zhang et al., 2015) used CiteSpace to map important references that lead to the development of SMKM trends, authors contributing greatly to this field and
hot topics of all the related articles. The development of SMKM study was analyzed according to the visualization of references and topics. Furthermore, the two most important groups – topics from SM and those from KM study were investigated, respectively to compare their development to show the fusion, separation and any other relationship. Finally, hottest trends and topics these years and the near future are discussed to provide the groundwork for future studies. In an article by (Makkizadeh et al., 2016) the scientific mapping approach was used to analyze the articles related to depression therapy. The findings indicated that drug therapy, psychology, anti-depressive agents, and treatment outcome are the most active research areas.

Thus far, there is no significant research with the aim to analyze the field of e-learning based on the approaches related to network analysis and scientific mapping. Conole et al. (2004) believed that a better articulation and mapping of different pedagogical processes, tools and techniques will provide a pedagogic approach, which is more reflexive and consistent with practitioners’ theoretical perspective on learning and teaching (Lee & Segev, 2012). In this study, an attempt was made to automatically build a domain knowledge map for e-learning, using text mining techniques. The experiments compared K-map learning to document learning, and found that K-map can identify the more important ideas. Barteit et al. (2020) conducted a systematic review of e-learning interventions for medical education in low- and middle-income countries. They pointed to the lack of high-quality research in this area, and emphasized on the need for more rigorous methods of evaluating the interventions, which will lead to a better assessment of medical e-learning and improved medical care in LMICs. In their study, Abrishami and Aliakbary (2019) utilized a machine learning tool to present a model for predicting the long-term citations of scientific papers based on the number of citations in the first few years of their publication. Based on their experiments, they claimed that their proposed method can provide more accurate predictions compared to the other advanced methods. In an experiment, Reffay and Chanier (2003) adopted the concepts of Social Network Analysis in a distance learning environment to measure cohesion in small groups of learners. They claimed that their proposed method is more appropriate than textual analysis of messages or the statistical distribution of learners’ contributions. Hamulic and Bijedic (2009) analyzed some features of a social network, including closeness and betweenness, to calculate knowledge sharing in a virtual learning community of students in the faculty of information technologies.

The evaluation results Begum and Nagahban (2019). showed that scientometrics is applied to assess the level of communication and efficiency in teaching-learning activities of individual learners. The evaluation results Wu and Wang (2012) showed that the dual mapping learning environment was perceived to be useful, innovative, and attractive. Overall, it was well received by the learners. there is a lack of studies on such cognitive support to facilitate learning through problem solving (Wu & Wang, 2010). Concept mapping has already been used within the Law School to identify the quality of learning of first year law students studying the law of contract (Hay & Proctor, 2015), but this would extend its use to encompass emotional aspects too. to identify relevant issues of pedagogic frailty or to consider the emotional impact of e-learning. The challenges of using concept mapping in this way include a potential reluctance on the part of the academy to accept or engage meaningfully with the role of emotion, linking to issues of pedagogic frailty and a desire to maintain a “safe” status quo without tackling the potentially difficult questions involved in utilising emotion constructively. (Jones, 2017). The analyses helped in identifying the various levels of knowledge actors, such as knowledge creators, brokers and users, in different knowledge area (Murale & Raju, 2014).
A review of literature indicates that so far, no independent research has been conducted on word co-occurrence and thematic clustering in the field of e-learning. The present research takes an analytical approach and aims to identify and map the research areas of this field, using the co-occurrence method. Scientific mapping and network analysis can help the policy makers to have a better knowledge of the structure of a research area, explore the research status of a field and develop plans to improve the quality and quantity of scientific output. For instance, the results of co-word analysis in a study conducted by (Makkizadeh et al., 2017) revealed the outstanding topics of the field under study. Identifying these priority research topics in any discipline can be useful for policy makers, since it helps them to learn more about the research status in order to make appropriate decisions for the promotion of seminal works in a discipline.

3. Method

This is a descriptive-survey research, using a scientometric approach. The statistical population of the study includes all the articles published by international researchers indexed in the Web of Science from 1995 to 2018. No sampling was performed in this research, and the entire data were collected and then analyzed. All data in this study were extracted directly from the Web of Science database, and scientometric software programs were indirectly used to perform calculations based on the extracted data.

The statistical population of the study includes all articles in the field of electronic learning, published by international researchers, and indexed in the Web of Science from 1995 to 2018. With respect to time and location constraints of the research, the following search query was used in the Advanced Search section of the Web of Science to extract data:

\[ \text{TS = (e-learning or electronic learning) AND PY = (1995-2018)} \]

The data were saved on the computer in Plain Text format. In the next stage, all the international papers published within the indicated time period were retrieved and extracted. At this point, the output files were merged into a single file, using TXT Collector software. In this way, they can be transferred to scientometric software systems. As for the next stage, the data were pre-processed, using Bibexcel software. Descriptive statistics (frequency, frequency percentage) were used in the form of tables and charts to analyze the data.

Software or tools can be used to conduct bibliometric or scientometric analysis (e.g., Jia et al., 2020). In this study, in order to map the scientific output in the field of e-learning, the software program WC10.exe was used to transform the format of the retrieved records. Then, they are mapped, using VOSviewer. Raw data were first imported into Bibexcel software. Using the WC tag, the co-occurrence network of categories was created and converted to net format. It was then mapped, using Pajek software.

4. Results and discussions

*RQ1*: How is the growth trend of scientific production in the field of E-learning at international levels?

From 1995 to 2018, the total number of e-learning related articles indexed in the WOS Citation Index was 29,575. The findings indicate that the number of e-learning related
articles in the world is increasing on an annual basis. In general, the growth trend of e-learning related articles in the WOS citation index has been on the rise.

Fig. 1. The growing trend of e-learning related scientific articles in WOS citation index (1995-2018)

The findings revealed that the global growth trend of e-learning articles in WOS citation index has been on the rise. Based on Fig 1, the number of articles published in this field has been increasing every year, and this growing trend has continued.

**RQ2: What are the emerging and prominent subject areas in the field of e-learning?**

The most widely used subject areas in the field of e-learning were specified, and high-frequency keywords such as e-learning, Internet, information technology, (health) care, management and blended learning are amongst the newly emerging and active areas in this field.

Fig. 2 is a co-word map of e-learning field during the years 1995-2018. In total, 96 keywords were identified in the years under review, and they were mapped, using VOS viewer software. Many clusters and links were identified in this map. As seen in Fig. 2, Internet, E-Learning and Blended Learning are the categories with hotspots, and more e-learning related articles are presented in these categories. The hotspots are the words and subject areas in which more articles have been written, and the red nodes indicate that in this decade, most of the research has been conducted in this area.

Fig. 3 shows the first five years of research, in which the clusters and links had not yet expanded. Some concepts like education, learning, and information are the categories with hotspots. The interesting point about this map is that the keywords “e-learning” and “Internet” are not visible here. One possible reason can be the limited scientific output during this period. Other prominent terms on the map are “neural networks” and “machine learning”. Their noticeable presence in the early years could reflect the role of computer technology and artificial intelligence in establishing the theoretical foundations of e-learning as a new field of research.
Fig. 2. A co-word map of the field of e-learning in the world (1995-2018)

Fig. 3. A co-word map of the field of e-learning in the world in the first 5 years (1995-2001)

Fig. 4 shows the co-word map of the last five years, and numerous clusters and links are identified on this map. As seen in this figure, the Internet, e-learning, machine learning, higher education, and technology are the categories with hotspots, and articles related to e-learning are to be found in these categories. Larger nodes and clusters
Knowledge Management & E-Learning, 12(3), 268–279

indicate that more research has been conducted on e-learning in this period, and accordingly, newer keywords have emerged. In Fig. 4, a more substantial change is seen in the new and emerging nodes and clusters. Some words like “health” and “care” indicate that e-learning has undergone a steady development in medical sciences and is still an active field. The term “management”, which is also noticeable in the main map (Fig. 2), is related to the field of educational management. This finding is in line with the result of previous review of the literature (Cheng et al., 2014).

Fig. 4. A co-word map of the field of e-learning in the world in the last 5 years (2014-2018)

**RQ3: What topics have high impact in the e-learning field?**

Fig. 5 depicts the co-occurrence network of keywords in the field of e-learning from 1995 to 2018, using the UCINET software. This network consists of 100 nodes (keywords) that are circular in shape and has over 26,000 edges marked with blue lines, which are divided into 7 clusters. It is a continuous network since the number of edges is greater than the number of nodes. The large number of its lines demonstrates the highly complex relations within an interconnected network.

In the science map of e-learning, which is based on the closeness centrality measure, the nodes with higher closeness centrality are shown in small circles. Since closeness centrality measures the distance between two words in the network, the nodes
with higher closeness have greater influence in the network. They play a more central role in the network, and are more accessible to the other nodes. Accordingly, based on Fig. 5, it can be inferred that such topics as Internet, distance learning, learning management system, MOOC and mobile learning have had the greatest impact in the above-mentioned map. In contrast, game-based learning, big data, and virtual labs appear in larger circles with the lowest centrality and the least impact.

**Fig. 5.** An overview of the co-occurrence network of keywords in the field of e-learning

**Fig. 6.** Co-occurrence network of descriptors based on the betweenness centrality measure

**RQ4: What topics serve as a bridge from one topic to another in the e-learning field?**

Being the structural property of a node, betweenness centrality indicates the importance of that node in terms of its location on the map and in terms of information transmission over the network. Fig. 6 shows an overview of the co-occurrence network of descriptors
based on the betweenness centrality measure. The keywords with the highest betweenness centrality measure including e-learning, education, MOOC, learning, online learning, blended learning, and learning management system. Therefore, these keywords fall in the middle of a large number of other nodes. They provide the major linking routes between other nodes and have the power to isolate or enhance communication.

5. Conclusion

The findings indicated a growing trend in the number of articles published in the field of e-learning, and this number is increasing annually from 1995 to 2018. In addition to the quantitative aspect, there are significant changes in the research topics in the period under study. A comparison of the maps from the first and the last five years of this period exhibits a shift from conceptual terms to more technical terms associated with e-learning.

High frequency keywords such as e-learning, Internet, IT, (health) care, management and blended learning are the newly emerging and active concepts in this field. Among them, e-learning and Internet are ranked first and second respectively as the most frequently used terms. The keywords “information technology”, “(health) care”, “management” and “blended learning” are the other high frequency terms representing the newly emerging concepts in this field.

Many more clusters and links were identified in this map. “Internet”, “E-Learning” and “Blended Learning” are the categories with hotspots, and more e-learning related articles are presented in these categories.

The field of E-learning was mapped based on the closeness centrality measure, and its findings revealed that topics such as internet, distance learning, learning management system, MOOC and mobile learning have the greatest impact on the map. In contrast, “game-based learning”, “big data”, and “virtual labs” appear in larger circles with lower closeness and less impact.

The keywords with the highest betweenness centrality include e-learning, education, MOOC, learning, online learning, blended learning, learning management system. They are the major linking routes between other nodes, which have the power to isolate or enhance communication.

Given the conceptual map formed in the field of e-learning, it can be concluded that the concepts on this map are not far apart from each other, and their close distance indicates that they are highly interconnected. In other words, there is a high density of subject areas or descriptors appearing in articles. This finding is in line with the findings of a research by Yang et al. (2011).

Furthermore, given the closeness and betweenness centrality measures in the provided map, it can be argued that the highest value (for both measures) is attributed to the subject areas of e-learning, learning management system and internet. This finding is in line with the findings of a research by Makkizadeh et al. (2016). In addition to being linked to many other research areas, these subject areas of learning are also important in terms of their positions on the map, so that the connections between many other subjects are only made possible through these subject areas.

Finally, it should be maintained that discovering the conceptual relationships among published articles in a scientific field is a highly complex process. Therefore, a profound outlook is required to understand the findings.
Author Statement

The authors declare that they have no conflict of interest.

Acknowledgements

N. Zarifsanaiey conducted the systematic literature review. M. B. Negahban wrote the draft manuscript. The manuscript was reviewed by M. B. Negahban. The authors wish to thank Mr. H. Argasi at the Research Consultation Center (RCC) of Shiraz University of Medical Sciences for his invaluable assistance in editing this manuscript. This work was supported by the Shiraz University of Medical Sciences [2019_18316].

ORCID

Mohammad Bagher Negahban https://orcid.org/0000-0001-6849-9555

Nahid Zarifsanaiey https://orcid.org/0000-0002-1297-8271

References


