

Original Research

Forecasting the Subject Trend of International Library and Information Science Research by 2030 Using the Deep Learning Approach

Farzaneh Ghanadi Nezhad

Ph.D. Candidate in Knowledge and Information Science, Shahid Chamran University of Ahvaz, Ahvaz, Iran, farzaneh.ghanadinezhad@gmail.com
ORCID iD: <https://orcid.org/0000-0002-5924-7169>

Farideh Osareh

Professor, Department of Knowledge and Information Science, Shahid Chamran University of Ahvaz, Ahvaz, Iran, f.osareh@scu.ac.ir, ORCID iD: <https://orcid.org/0000-0001-6691-0339>

Mohammad Reza Ghane

Associate Prof. of LIS, Department of Evaluation and Collection Development, Regional Information Center for Science & Technology (RICeST), Shiraz, Iran.
Corresponding Author: ghane@ricest.ac.ir, ORCID iD: <https://orcid.org/0000-0002-1832-7904>

Received: 09 August 2021

Accepted: 06 October 2021

Abstract

This study seeks to forecast the subject trend of library and information science research until 2030 based on modeling previous research topics in this field, which has been done with a text mining and in-depth learning approach. After pre-processing and thematic classification of the studies, deep neural network algorithms were used to model previous studies and forecast future topics. The study population included 90,311 journal articles in library and information science publications indexed on the Web of Science website from 1945-2020. All research processes were implemented in the Python programming language. The findings showed that the largest number of studies in the future would be related to Internet and web studies, and the growth rate of these topics will be higher in the future. However, topics related to libraries and their work processes and other traditional disciplines such as theoretical foundations will have a lower growth rate in library and information science studies. As a result, knowledge of important future issues, while helping to plan for future research, can identify study gaps and investment opportunities in the R&D sector, thereby assisting researchers, universities, and relevant research institutes in selecting projects intelligently.

Keywords: Subject Forecasting, Research Trend, Subject Trend, Future Topics, Library and Information Science, Deep Learning.

Introduction

In the era of rapid developments, especially digital developments, the need for systematic studies and reflections on the future is felt in library and information science to adapt to modern needs. Planning and preparation to face future challenges and maintain status and survival in

these competitive conditions require such systemic studies so that this field can track and follow developments and forecast future changes. Otherwise, the field will be left out of the competition and will lose its position in the future. As Godet (2006) points out, with future studies, in addition to analyzing past trends, it is possible to invent and evaluate possible, probable, and desirable futures. A review of studies shows that researchers in various fields have realized the importance and complexity of the problem of research future and forecasting the subject trend of scientific productions and have always tried to forecast future publishing trends from different perspectives using different methods. Identifying emerging issues is one of the most critical goals used to forecast the subject trend of scientific productions. Emerging issues analysis is a term in future studies and strategic management and policymaking that addresses the process of identifying topics that were not important or influential in the past but may be influential in the future (Tavakoli, Babakirad & Ghoroonch, 2016).

Different methods have been used to determine emerging topics and forecast future research trends in these studies. One of the common approaches in this field is to use experts' opinions and mental judgments (Singh, Hu & Roehl, 2007; Salatino, 2015; Zhang, Zhang, Chen, Porter, Zhu & Lua, 2016; Li, Xie, Daimb & Huang, 2019). However, it should be noted that forecasting the future may be biased by experts or that those that participate in the survey may not have sufficient expertise. It may also be challenging to calculate the level of confidence in the forecasts. On the other hand, due to the continuous increase of scientific productions and their rapid developments, it is difficult for professionals to keep up with the rapid changes in research (Salatino, 2015; Chen, Wang, Li & Sun, 2018; Abuhaya, Nigmatie & Kovalchuk, 2018). The changes make the use of approaches based on expert opinions less practical. More attention is currently being paid to automated and scalable methods to calculate the future trend of scientific production. Forecasting the subject trend of research in a completely automatic manner has also been considered. De Solla Price (1965) first introduced the concept of active research fields as growing fields to show the future direction of topics. Later, co-citation analysis was proposed by Small (1973, cited in Small, 2006) to identify emerging topics for future research. During 1974-2015, the analysis of citation networks through co-citation analysis, bibliographic pairs, and a direct citation was mainly used to identify emerging subject areas in research (Small, 2006; He, Chen, Pei, Qiu, Mitra & Giles, 2009; Small, Boyack & Klavans, 2014). It can be noted that in this period, citation-based approaches based prevail in this area (Xu, Hao, An, Pang & Li, 2020). During and after this period, some studies have used vocabulary-based approaches such as keyword analysis and co-word analysis to study the future direction of research (Chen, 2006; Guo, Weingart & Borner, 2011; Small et al., 2014; Wu & Ren, 2018; Hoz-Correa, Munoz-Leiva & Bakucz, 2018; Lara-Rodríguez, Rojas-Contreras & Oliva, 2019; Pestana, Sánchez & Moutinho, 2019).

It should be noted that the use of quantitative analysis methods such as bibliometrics, scientometrics, and informatics to determine the direction of future research has some limitations, such as time delays resulting from the use of citation-based methods (it takes time for research to be cited). This causes bibliometric and co-citation analyses to be based on older sources, and their results cannot show recent subject trends and the future direction of research. Moreover, the large number of citations may not necessarily indicate the higher quality of the research works and topics. This is because people can cite more of their research or their colleagues or the journals they choose to publish their findings. Such problems and the increasing volume of textual resources have led to limitations in applying approaches based on

vocabulary analysis, such as the homologous method, for future studies of scientific products. These issues have led researchers to think of alternative or complementary solutions and methods for identifying emerging fields of study and future research perspectives (Abuhaya et al., 2018).

Since 2015, there have been developments in identifying emerging research topics and new approaches to forecasting the trend of scientific production. In this regard, Joung & Kim (2017) proposed a keyword-based model for analyzing data of patent databases to track emerging technologies using links between keywords and the term frequency-inverse document (TF-IDF) method. In another study, Porter, Garner, Carley and Newman (2018) identified and presented indicators for identifying hot topics to identify emerging research areas and improve scientometric indicators. During this period, along with the development of large-scale word processing technologies, machine learning methods were gradually used to identify emerging areas and forecast the research process because of their high power in analyzing large data (Lee, Kwon, Kim & Kwon, 2018). Use of ontology and keyword relationships (Salatino, 2015; Balili, Segev & Lee, 2017), and data mining, text mining, and machine learning algorithms (Singh et al., 2007; Wu, Shiu, Lee & Su, 2010; Budi, Aji & Widodo, 2013; Balili et al., 2017; Bildosola, Gonzalez & Moral, 2017; Bai, 2018; Sohrabi & Khalili Jafarabad, 2018; Chen et al., 2018; Li et al., 2019; Choi, Park & Lee, 2021; Cho, Lee & Sohn, 2021) are other strategies that can help identify emerging topics and future research priorities. As Santo, Coelho, Santos & Filho (2006) point out, in science and technology forecasting, currently more efforts are being made to use machine learning approaches, deep learning, and modeling techniques to find patterns and precise relationships between data and inferential rules that make it possible to identify future trends rather than traditional scientometric methods.

Studying the future of library and information science is done in different dimensions. One of the essential axes is studying research needs and future study fields. Because of the interdisciplinary nature of library and information science, attention to future subject areas in this field is of particular importance. The primary studies regarding subject trends in library and information science research analyze articles published in journals at different periods (Tuomaala, Järvelin & Vakkari, 2014; Luo & McKinney, 2015; Liu & Yang 2019). They dealt with the subject analysis of dissertations (Zong, Shen, Yuan, Hu, Hou & Deng, 2013; Anna, Mannan, Srirahayu & Mutia, 2018), conference papers (Garner, Davidson & Williams, 2008), the structure and field mapping (Zins, 2007; Figuerola, Marco & Pinto, 2017; Han, 2020), content analysis of a collection of field texts including articles, dissertations, etc. (Xiao, Zhang & Li, 2015), and content analysis of the content of other texts in the field, including research priorities (Sarah & Pete, 2004). These studies have mainly analyzed the past and present trends of scientific outputs in this field using content analysis and scientometric approaches. Attention to the subject trends of future research in this field has been somewhat neglected.

The present study seeks to forecast future research in library and information science based on past research in this field. Considering the need to adapt and keep pace with future developments, plan for effective confrontation with possible future problems and challenges, and solve them. Having a picture of future subject areas in this field will guide researchers to identify the most critical issues and needs and appropriate action to select the topic for future research. However, suppose the future needs are not addressed. In that case, a significant part of research resources and credits will be spent on topics that do not have the necessary priority and importance. We cannot expect to use their results to solve problems and develop this field.

Studies have been conducted in library and information science to determine the future direction of research and identify emerging topics. In these studies, scientometric and co-word approaches have been used. Studies have shown topics such as digital libraries, information retrieval, knowledge management, information and communication technology, Internet, web technology, semantic web, social media, search engines, metadata, ontology, emotion analysis, natural language processing, machine learning, data mining, text mining, and web mining are new and emerging fields in the field (Summers, Oppenheim, Meadows, McKnight & Kinnell, 1999; Kumar Sinha, 2016; Soheili, Khasseh & koranian, 2018; Mansourkiaie, Babalhavaegi, Nooshinfard & Soheili 2019; Abdollahzadeh, 2019; Baghmohammad, Mansouri & Cheashmehsohrabi, 2020; Taşkın, 2021). Given the importance of conducting such studies for future policies and planning, we aimed to use the deep learning approach to forecast the trend of international research topics in the field of library and information science to take a step towards identifying the needs of future research.

Research Methodology

The present study was conducted with a text mining and deep learning approach. After pre-processing and thematic classification of the studies, deep neural network algorithms were used to model the subject trend of past studies and forecast future trends based on past data models. The study population included articles of all library and information science journals that were indexed on the Web of Science (WoS) from 1945-2020 and is listed under the heading "Information Science and Library Science" in the Journal Citation Reports (JCR). Thus, 89 publications containing 90,311 articles were retrieved. The publication name (SO) field retrieved all articles in the Advanced Search section. Records retrieved in the document type section were limited to articles, English language, and the subject category section to information science and library science. The most critical executive steps of the present study were as follows:

Selecting the subject field to be analyzed

To perform the analysis of the present study, the titles of the articles were used. First, the bibliographic information of articles, including title and year of publication, were extracted from WoS database. Then a database was created based on the information stored in Excel 2016 software.

Selecting the subject classification plan of scientific products

The classifications presented in specialized databases have often been used in previous studies to classify studies in this field. However, it should be noted that each of these classification schemes alone cannot cover the various aspects of library and information science and novel and interdisciplinary subject areas. Moreover, research fields in this area are influenced by various social, cultural, and technological developments in the community and within the field. These developments indicate the need to review existing classification schemes. It was necessary to set up a system that would cover all the subject areas of research in this field, including traditional and modern topics, in other words, all the subject areas of the past 76 years and the outline of future topics in the field. Therefore, by studying and combining the existing classification schemes and consulting experts in this field, researcher-made lists were prepared in 31 subject areas and became the basis of analysis.

Data Preparation and Pre-processing

In order to convert the titles of the articles into a standard and structured format and standardize the words and phrases derived from them, data pre-processing operations were performed. These operations include removing stop words, stemming (in a structured format) and lemmatization, deleting duplicate data, outliers, and missing values, and extracting concurrent words and vocabulary using n-gram methods (ngram_range=[1, 3]). Finally, the TF-IDF method was used to select valuable features and keywords, and the documents were displayed as a vector space model (VSM), known as the Bag of Words (BOW), and the term-document matrix was formed. We normalized data using the Min-Max Scaler algorithm. The matrix dimensions were reduced with the PCA algorithm, English documents were displayed as a matrix of terms in the dimensions of 8438*77161. The occurrences of each term in a document were recognizable in the title of the articles. Furthermore, the weight of that word was identifiable.

Subject classification of the studies

In this research, the automatic text classification method, one of the text mining applications, was used to classify studies in this field. The results showed that among the models used for classification, the model obtained from the logistic regression algorithm for classification of researches had a precision, accuracy, recall, and F1 score above 90% and, as a result, had a better performance in document classification.

Use deep neural networks to forecast the future

In the present study, deep neural networks were used to forecast the subject trend of the studies. The network architecture used in this study is long short-term memory (LSTM). LSTM is a recurrent neural network designed to store, process, and access information better than conventional neural networks. These networks have a return loop that allows the information they have already obtained not to be lost and to remain in the network. However, conventional neural networks used by machine learning professionals have not behaved in this way, which is a significant drawback for these networks (Olah, 2015).

It should be noted that recurrent networks are not capable of long-term dependency learning and do not have the power to recall and use the information they have learned in the distant past. In other words, they cannot use information from the distant past (Hochreiter, 1991; Bengio, Simard & Frasconi, 1994). In this regard, LSTM networks were designed to solve the problem of long-term dependency.

Design and implementation of the deep network model

Before implementing the neural network model, the data must be transformed into columns, acceptable for neural network model inputs. This operation was performed using the reshape function in the NumPy library. Then, the input data was normalized using the Min-Max Scaler function. The most important part of designing a deep network model is defining its layers. In order to place the deep network layers, the sequential model, which is the simplest and most basic type, was used. The layers are placed one after the other.

Furthermore, they are defined as a sequential and linear set of layers. Each layer is associated only with the next layer. In fact, in this model, different layers are added sequentially.

The architecture of the designed model for the deep network in this study included a dense input layer (LSTM), two hidden dense layers, two dropout layers, and a dense output layer, the functions of which will be described in the following section.

Input layer: The main layer and the input of the deep network that receives the data is of the LSTM type. The network input data was then converted into columns. The information in each column is related to 31 subject axes, which includes 76 data related to the years 1945 to 2020 for each subject axis. The problem parameters include $X_1, X_2, X_3, \dots, X_{76}$, each representing the data for each year. For example, X_1 is the data for 1945, and X_{76} is the data for 2020. To define each layer, we must first determine the number of neurons in that layer. The number of these neurons is a parameter obtained by trial and error. Finally, six neurons were identified to be suitable for this layer. The hyperbolic tangent function (Tanh) was used in the input layer. The reason for choosing the Tanh function was that in the model's design by the software itself, it was placed by default and had a minor error during the test of different functions.

Hidden dense layer: A dense layer with 300 neurons was added as a hidden layer to improve the function of the neural network. The rectified linear activation function (ReLU) was used in this layer, the most common and most used function in the hidden layer.

Dropout layer: This layer improves network performance in deep networks. The dropout layer specifies how many percent of network neurons are accidentally turned off during the training process or have zero input to reduce repetitive learning between neurons, speed up network learning, and prevent overexertion. In the present neural network model, this layer determined that 0.3 neurons were excluded during the training process.

Hidden dense layer: In the network designed for the present study, another hidden dense layer with 100 neurons was added, and in this layer, the ReLU was used.

Dropout layer: A dropout layer with a value of 0.1 was also placed between the hidden layer and the output layer. This means that 0.1 neurons were left out for the model training process.

Output layer: A dense layer with a neuron was placed in the output section.

Compiling the model: After pre-processing and normalizing the data and defining the model architecture, the constructed model was compiled using the *model.compile* command. In this section, the loss function was used to show the network performance of data training. At this stage, the ADAM optimization algorithm, which is one of the most popular optimization algorithms and has high performance in systems with low memory, was used to improve the network training process designed.

Model fit: Finally, the model was assessed using the *model.fit* command, and the model was executed on the data. This step determines how many times it is necessary to apply the training to the data; each is called an epoch. It should be noted that the model will have some errors each time the training process is repeated, but after a while, this number of errors becomes uniform. At this point, the model is said to have converged around this level of repetition of the training process. In the present study, the repetition rate of network training was based on this, and the model converged around 100 epochs. Another parameter that is specified in this step is batch size. The function of this parameter is such that the model calculates the average of errors after reviewing some data. Then optimizes and updates the network based on this average. The batch size in the present research model was set to 30.

The research data was divided into two categories: educational and experimental to measure

the efficiency of the designed model. Consequently, 80% of the documents (1945 to 2005) were used to train the model, and 20% (2006 to 2020) were excluded for model testing. Then, the implemented model was applied to the educational data. The learning rate of the model and its performance were evaluated to forecast the test data. In order to estimate the error rate, we compared the difference between the actual values with the trained values. Moreover, the difference between the actual values and the tested data values was compared. Also, the error rate in learning the model and forecasting it was calculated. In this regard, the mean square error (MSE) method was used.

The programming language used in this study was Python. Pandas, Numpy, Hazm, and Nltk software packages were used for data entry, processing, matrix, word formation, classification, and testing, and Numpy, Pandas, TensorFlow, and Keras packages were used to build (implement) the deep neural network and forecast the future.

Results

The subject trend of research in library and information science research in the future over the past 76 years was studied. Table 1 presents the distribution of research dedicated to each of the subject areas of the field during the previous years.

Table 1

The Subject Trend of Library and Information Science Research on the Web of Science Website during the Years 1945-2020

Research fields	Frequency	Perentage
Information and Communications Technology	7075	9.17
Internet and web studies	6535	8.47
Libraries, archives, and information centers	4314	5.59
Research, study, and publishing	3907	5.06
Data saving and information retrieve	3092	4.01
Scientometrics and Informatics	2944	3.82
Information systems	2889	3.74
Information and knowledge-based services	2883	3.74
Databases	2759	3.58
Teaching and learning	2755	3.57
Information resources management	2617	3.39
<i>Information and knowledge organization</i>	2561	3.32
User studies	2427	3.15
Scientific and information communication	2370	3.07
knowledge management	2327	3.02
Information Behaviors	2277	2.95
Profession, employment, and entrepreneurship	2239	2.90
Librarians and information specialists	1940	2.51
Electronic, digital, and virtual libraries	1934	2.51
Search engines	1902	2.46
Economics and information marketing	1832	2.37
Ethics and information security	1832	2.37
Information processing and management	1668	2.16
Information and digital literacy	1530	1.98

Research fields	Frequency	Percentage
Data mining and knowledge discovery	1480	1.92
Artificial intelligence	1426	1.85
Information and knowledge-based society	1243	1.61
Future studies in library and information science	1221	1.58
Theoretical foundations of library and information science	1171	1.52
Semantic web and ontology	1071	1.39
Organizational management	940	1.22

As shown in Table 1, at the international level, most articles were in the fields of information and communication technology (9.17%), Internet and web studies (8.47%), Libraries, archives and information centers (5.59%), Research, study and publication (5.06%), and data saving and information retrieve (4.01%). In contrast, the fields of organizational management (1.22%), semantic web and ontology (1.39%), theoretical foundations (1.52%), and future studies (1.58%) have the lowest number of articles.

In order to forecast the subject trend of library and information science research, in the first place, the efficiency of the designed deep network model was tested. As mentioned in the present study, the loss function was used to show the network performance in data training. The results are shown in Figure 1 for training data related to the first axis (theoretical foundations).

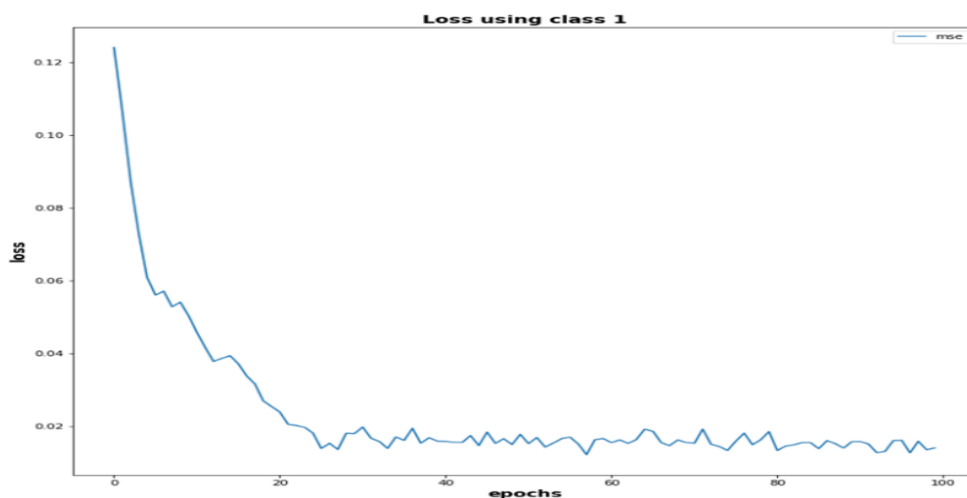


Figure 1: Results of the Error Rate during the Deep Network Model Training Process for the Theoretical Foundations

Figure 1 shows that during the 100 times performing the model training process, the error rate is reduced, and network performance is improved. The model's performance concerning other axes was the same, but we have not brought their volume in this study due to their high volume. After executing the implemented model on educational data, the learning rate of the model and its performance to forecast the test data were evaluated. In order to estimate the error rate, the difference between the actual values and the trained values and the difference between the actual values and the values of the tested data were compared, and the error rate in learning the model and forecasting it was calculated using the MSE method. The results of comparing

the actual values of the theoretical basis with the trained values and the values forecasted by the model are shown in Figure 2.

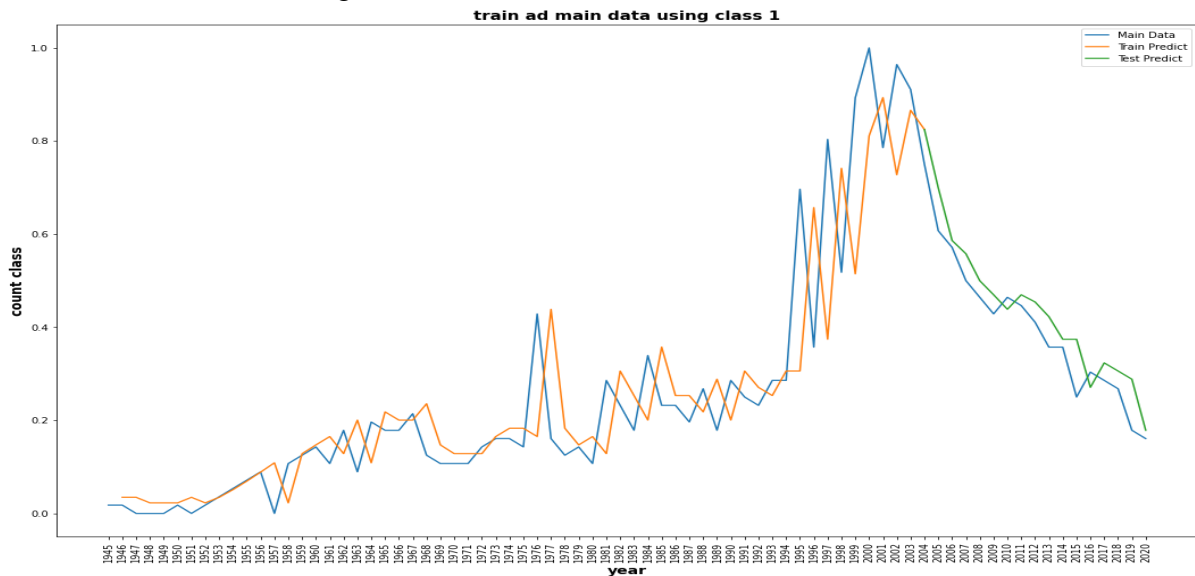


Figure 2: Comparison of Actual Values Based on Theoretical Foundations with Trained Values and Values Forecasted by the Deep Network Model

Figure 2 shows that the data trained and forecasted by the model are close to the actual data in terms of theoretical foundations, which means that the designed model is well trained and has a good performance in evaluating the data. In the following, the data obtained from the calculation of the MSE criterion, i.e., the error score in the training and test section of the model, are listed in Table 2.

Table 2

Calculating the Mean Square Error in the Training and Testing of the Forecasting Model at The International Level

Row	Research fields	Mean square error	
		Education mean	Test Mean
1	Theoretical foundations of library and information science	0.77	0.28
2	Libraries, archives and information centers	0.49	0.03
3	Organizational management	0.02	0.57
4	Electronic, digital and virtual libraries	0.03	0.14
5	Profession, employment and entrepreneurship	0.04	0.72
6	Teaching and learning	0.12	2.46
7	Research, study and publication	0.22	2.47
8	Information and communications technology	0.03	3.45
9	Internet and web studies	0.03	2.43
10	Information processing and management	0.05	1.88
11	knowledge management	0.02	0.74
12	Information and digital literacy	0.06	0.09
13	Future studies in library and information science	0.11	2.62
14	Information resources management	1.42	1.07

Row	Research fields	Mean square error	
		Education mean	Test Mean
15	<i>Information and knowledge organization</i>	0.57	0.47
16	Data saving and information retrieve	1.27	2.46
17	Information and knowledge-based society	0.15	1.49
18	Semantic web and ontology	0.01	0.90
19	Economics and information marketing	0.04	0.60
20	Information Behaviors	0.03	0.26
21	Artificial intelligence	0.02	1.25
22	Data mining and knowledge discovery	0.03	2.62
23	Scientometrics and informatics	0.01	2.32
24	Information systems	0.62	0.54
25	Search engines	0.009	1.45
26	User studies	0.12	1.57
27	Librarians and information specialists	0.10	1.18
28	Information and knowledge-based services	0.26	0.63
29	Scientific and information communication	0.16	1.95
30	Databases	0.02	1.03
31	Ethics and information security	0.21	1.78

Table 2 shows that the error obtained for training and testing the model in all axes was very small. The higher error rate in the information and communications technology axis, among other axes, has been due to the more irregular trend and sudden jump in these issues in recent years.

After ensuring the accuracy of the deep network design forecasting, the model was re-trained, and forecasting the research topics for the next ten years was performed. Network implementation and future forecasting were done using a Python program designed for this purpose. According to data columns, 76 data from 1945 to 2020 were used to form a network and forecast the future. In this way, data from the last 76 years are given to the network, and the network is trained accordingly. Based on past trends, the data criteria and rules are then extracted and forecasted for the next decade. In other words, the deep network acts as a system that uses the value of the input data to calculate the output data. Each of the connections between data in different layers has its weight. The network learns the pattern between input and output variables by adjusting these weights during the training phase. The final output of the deep network model is to forecast the number of studies in each of the 31 areas of study in the library and information science for the next decade (i.e., 2021 to 2030, which is shown in Table 3 with data for the last six years).

Table 3

The Subject Trend of Scientific Productions in the Field of Library And Information Science at the International Level in the Last 6 Years and Forecasting this Trend in the Next 10 Years

year of publication Research fields	Number of research titles per year															
	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
Theoretical foundations of library and information science	14	17	16	15	10	12	13	16	14	18	17	14	13	11	9	9
Libraries, archives and information centers	62	60	58	56	51	50	47	52	49	48	45	42	38	36	33	29
Organizational management	34	31	46	41	48	34	37	27	34	31	29	38	38	31	36	37
Electronic, digital and virtual libraries	103	105	108	119	128	132	132	138	141	140	143	148	150	154	158	161
Profession, employment and entrepreneurship	94	95	98	113	133	129	130	137	140	144	147	150	153	156	159	162
Teaching and learning	119	120	121	126	130	138	141	143	146	146	147	148	149	152	157	158
Research, study and publication	175	188	205	208	212	223	226	228	231	234	236	237	238	240	241	241
Information and Communications Technology	351	360	361	412	418	433	445	454	461	468	474	479	484	488	491	494
Internet and web studies	449	461	486	510	517	530	549	563	573	578	583	586	587	895	915	597
Information processing and management	77	91	102	113	129	130	133	131	133	135	136	138	139	143	141	142
knowledge management	157	144	149	153	164	166	169	171	174	173	175	177	179	182	185	189
Information and digital literacy	94	97	99	100	114	109	114	116	116	117	119	120	122	123	123	125
Future studies in library and information science	71	71	80	85	91	104	109	114	116	119	121	121	123	124	126	127
Information resources management	40	39	37	37	36	33	32	30	29	27	26	25	23	22	20	19
Information and knowledge organization	49	47	43	41	41	40	39	38	37	36	38	36	35	33	31	30
Data saving and information retrieve	85	85	92	94	108	119	123	129	129	139	144	149	156	158	166	170

Research fields	Number of research titles per year															
	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
Information and knowledge-based society	54	46	59	46	50	56	61	59	67	66	75	74	84	83	94	95
Semantic web and ontology	67	77	82	96	101	117	128	142	147	153	158	163	168	172	177	178
Economics and information marketing	113	126	139	147	156	169	175	179	183	186	189	192	194	196	198	199
Information Behaviors	142	149	163	170	186	188	194	199	207	213	218	222	224	227	229	233
Artificial intelligence	64	78	83	94	102	115	118	129	135	139	141	152	160	169	179	181
Data mining and knowledge discovery	92	101	106	120	144	152	158	160	167	172	178	182	186	192	195	198
Scientometrics and Informatics	182	178	152	139	119	117	115	114	111	115	109	108	105	103	100	98
Information systems	81	81	84	81	88	78	81	74	73	75	75	74	76	77	76	78
Search engines	121	127	141	152	154	160	175	182	188	189	192	195	197	199	202	207
User studies	104	104	117	118	136	138	142	147	148	152	155	159	161	165	167	170
Librarians and information specialists	47	50	62	69	78	78	90	93	94	97	100	108	114	117	124	126
Information and knowledge-based services	91	90	89	82	89	90	87	90	91	94	93	94	95	91	95	96
Scientific and information communication	83	117	112	111	103	113	109	111	115	114	119	125	121	125	126	128
Databases	113	120	129	139	153	154	157	158	161	165	167	176	177	180	184	186
Ethics and information security	76	68	76	77	93	89	88	93	95	96	97	98	99	100	101	102

According to the results obtained from Table 3, the topics that have had an increasing trend in recent years will continue to grow in the following periods. The topics that have had a decreasing trend in the following years will also decline. Thus, the areas of theoretical foundations; Libraries, archives, and information centers; information resource management; information and knowledge organization; and scientometrics and informatics, which have been declining in recent years, are expected to continue to decrease next decade. Other topics that have faced an upward growth trend in recent years will continue their upward trend in the next period with slight fluctuations. Figure 3 shows the forecast of the number of studies dedicated to the library and information science fields at the international level for the next ten years, respectively.

As shown in Figure 3, in the future, the most significant number of studies at the international level will be in the fields of Internet and web studies, information and communication technology, research, study and publishing, and information behaviors. Areas like theoretical foundations, information resource management, organizational management, and information and knowledge organization will have the least research. It is worth considering comparing the past subject trend of research in this field with their future trend. In the past, libraries, archives, and information centers have been among the areas with the highest amount of research, while the results show that there will be less research in the future. The domains of semantic web and ontology and future studies, which have received less attention from researchers in the past, will have more fans among researchers in the future. Another issue is forecasting the growth percentage of research topics in the next decade, shown in Figure 4.

As shown, the growth trend of research in many library and information science studies fields is increasing internationally in the next decade. However, considerable growth is seen in research in Internet and web studies, artificial intelligence, semantic web and ontology, information and communication technology, and data saving and information retrieval. According to the forecast results, the topics of information resource management, *Information and knowledge organization*, scientometrics and Informatics, libraries, archives and information centers, and theoretical foundations will experience a much slower growth trend.

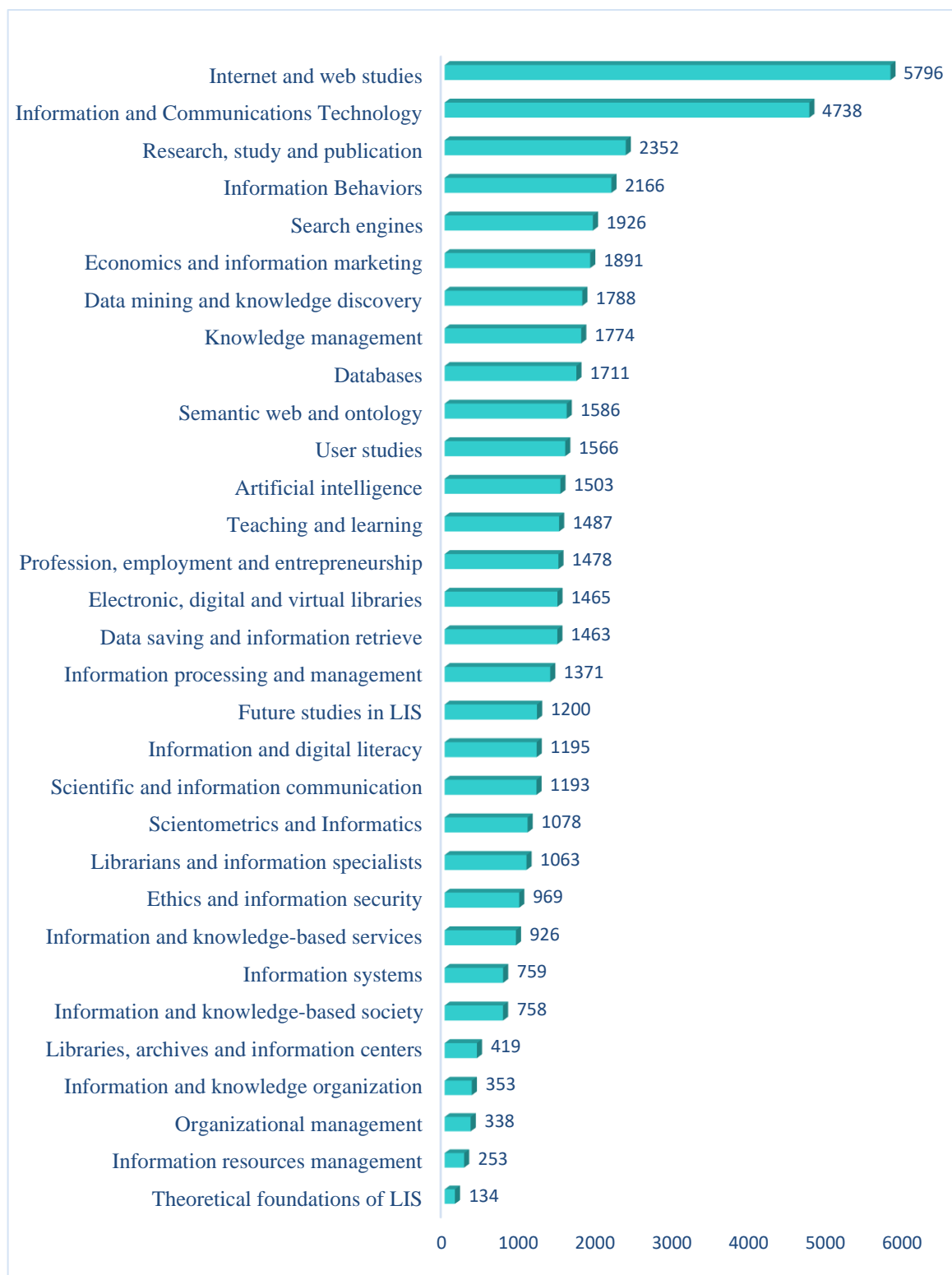


Figure 3: Forecasting the number of studies dedicated to each of the fields of library and information science studies at the international level for the next 10 years in order of priority

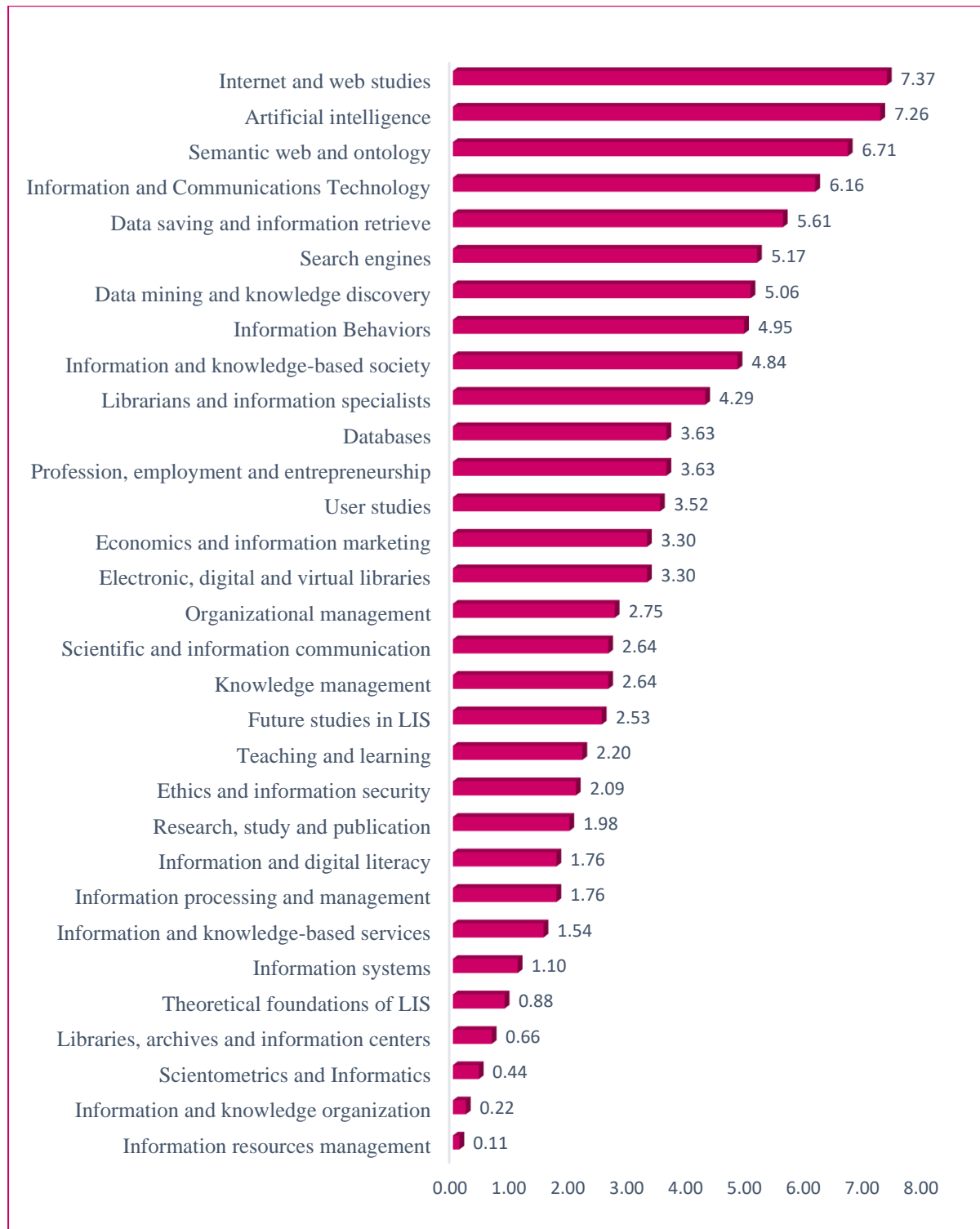


Figure 4: Forecasting the growth rate of studies dedicated to each of the fields of library and information science studies at the international level for the next 10 years in order of priority

Discussion

The continuation of comprehensive developments in the future and their impact on the scientific field reveals the need for systematic studies on future knowledge. One of the most important and influential aspects of library and information science affected by future developments is studying research topics. In this regard, the subject trend of future research in

this field was forecasted at the international level in the present study. Findings from the forecasting of the subjects of this field at the international level indicated that the highest number of studies in this field in the future would be in the fields of Internet and web studies, information and communications technology, research, study and publication, and information behaviors. Considering the results of the growth percentage of study axes at the international level, we found that the growth trend of research in many subject areas of this field is increasing in the next decade. However, the more rapid growth of Internet and web studies, artificial intelligence, semantic web and ontology, information and communications technology, and the data saving and information retrieve shows that these findings can be related to the adaptation of these areas to future developments and therefore the urgent need of library and information science to use the findings of these areas. These findings could also be related to the effects of the achievements and emerging technologies of the Fourth Industrial Revolution, such as artificial intelligence, robotics, blockchain, and the Internet that are rapidly taking shape. According to Schwab (2017) in *The Fourth Industrial Revolution* book, experts forecast that the developments resulting from this revolution will reach their peak by 2030.

The development of new information and communication technologies and the explosive growth of information have created challenges in the effective search and retrieval of large volumes of information and the analysis and interpretation of retrieved information. Therefore, experts in information retrieval have conducted studies to provide methods and solutions to solve these problems. Hence, the most important areas of research in library and information science include topics such as "smart agents", "ranking algorithms", "cluster analysis", "data mining and web mining", "graph web and semantic web algorithms" and "Internet studies and Web" (Carlson, 2003). Since the 1990s, due to the interdisciplinarity nature of library and information science, the subjects related to technology and the web have experienced more growth. Notably, this phenomenon is rooted in the growth of information and communication technologies and the emergence of the Internet.

Along with the explosive growth of information and the role of information and knowledge in economics, the new era affected the subject trends. It seems that in the future, the field of library and information science will have strong relationships with other areas at the international level and be interdisciplinary, especially regarding disciplines such as information technology, computer science, artificial intelligence, management, and communication (Mostafavi, Osareh & Tavakolizadeh-Ravari, 2018; Tang, 2004; Chang & Huang, 2011; Hessey & Willett, 2012; Dora & Kumar, 2019; Han, 2020). Therefore, the fluid nature of these fields and the rapid evolution of tools and technologies cause constant research focus changes (Saracevic, 1999, as cited in Han, 2020).

In confirmation of the present study's findings, we can refer to the results of some previous studies. Mansourkiaie et al. (2019) concluded that from the perspective of library and information science experts, areas related to information retrieval would be among the top research priorities in this field in scientific social networks in the next ten years. Moreover, issues related to policymaking in science and technology, training for information work, and information-seeking behavior models will be a priority for future research in this field. According to Abdollahzadeh (2019), experts in this field in recent years have used the facilities of information and communication technologies and, with the help of experts in other fields, have tried to remove barriers to information retrieval. With this description, it can be argued that information retrieval will be one of the most critical research issues in library and

information science so far and in the future, which helps describe the challenges facing users and find solutions to them. The findings of Ghanadinezhad and Heidari's (2018) study showed that from the perspective of library and information science experts, the most important topics for future research in this field were applied topics, topics based on the future of the field, and topics related to information and communication technologies. According to Summers et al. (1999) estimations, intending to chart the future of information science for the next ten years, the essential topics in this field in 2010 included digital libraries, information retrieval, and knowledge management. Also, a study analyzing the content of the call for seminars, conferences, and library and information science workshops in India and around the world from 2012 to 2014 found that applied topics based on information and communication technology, the Internet, web technology, and the semantic web, user studies, digital technologies, and social media represent the perspective of research in this field in the future scenario of India (Kumar Sinha, 2016). In addition, Abdollahzadeh (2019) concluded that emotion analysis, natural language processing, machine learning, data mining, text mining, web mining, social media, and metadata are among the emerging topics in library and information science are expected to grow faster in the coming years. The results of studies that have examined the subject trend of library and information science research in recent decades show that information retrieval has been one of the most critical subject clusters in this field in all periods. Nevertheless, research in recent years has shifted to machine learning, text extraction, topic modeling, emotion analysis, and text processing based on models or algorithms (Han, 2020; Taşkın, 2021).

The present study results showed that future researchers would be less desired by theoretical foundations, libraries, archives and information centers, information resource management, organizational management, and information and knowledge organization. In addition, the growth rate of these topics will be lower in the future. The downward trend in these areas may be since these areas of study have evolved. In other words, these subject areas are still relevant in the field, only under the influence of changes in society and the impact on this field, they have been proposed with a new title or concept. For example, library studies have led to electronic, digital, virtual, and intelligent library studies. Acquisition or information resource management research is more commonly referred to as information management. Traditional topics of information organization such as cataloging and classification have given way to cyberspace organization, metadata, automated classification, and automated indexing. In this regard, Mostafavi et al. (2018) acknowledged that in the past, traditional tools such as thesauruses, Subject headings, and classifications were used as tools for knowledge organization in information centers. Today, however, the Internet, the semantic web, ontologies, metadata, knowledge organization systems, electronic thesauruses, electronic databases, and multimedia tools have revolutionized how information is organized.

A noteworthy point among the present study's findings is that studies in the fields of scientometrics and Informatics, which have experienced significant growth in recent years, will be studied less in the future and will experience a slower growth trend. In justifying the downward trend of these studies, it can be pointed out that the evolution of information carriers and the increasing volume of textual, audio, video, and multimedia resources and the problems related to their analysis have led to limitations in applying traditional scientometric-based approaches such as the co-word method (Abuhaya et al., 2018). Along with the growth of information in various formats and media and the development of large-scale word processing

technologies, artificial intelligence approaches, machine learning, deep learning, and data mining and text mining methods due to their high power in processing and analyzing big data (Lee et al., 2018), modeling techniques (Santo et al., 2006), the use of ontology and keyword connections (Salatino, 2015; Balili et al., 2017) and semantic analysis of words (Li et al., 2019) have replaced scientometric methods. It is evident that in recent years and even in the future, the number of studies conducted with traditional scientometric approaches will decrease.

Conclusion

The occurrence of rapid developments in various aspects of life as a result of technologies and information and communication networks, the Internet and social media, and the continuation of these developments in the future and their impact on the field of library and information science, shows the need for future studies in this field in various dimensions, including the analysis of future research topics. Despite the rapid developments in research topics on library and information science and the problems associated with forecasting these topics in the future, if we understand the true nature and mission of the library and information science, we can expect to understand the future direction of research in this field more efficiently. Awareness of critical future issues, while helping to plan for future research, can identify study gaps and investment opportunities in R&D, thereby assisting research centers in selecting projects intelligently. With such an approach, it is possible to avoid repeating research topics with different titles and focus on new and practical topics with high potential for community development. Suppose the planners and decision-makers, especially in education and research, ignore fundamental issues and the final mission of this field in general. In that case, this field may lose its place in the scientific community in the future.

References

- Abdollahzadeh, P. (2019). *Mapping research topics of Library and Information Sciences based on co-word analysis*. Master Thesis, Tabriz University of Medical Science.
- Abuhaya, T., Nigmatie, Y. & Kovalchuk, S. (2018). Towards predicting trend of scientific research topics using topic modeling. *Procedia Computer Science*, 136, 304-310. <https://doi.org/10.1016/j.procs.2018.08.284>
- Anna, NE. V., Mannan, E. F., Srirahayu, D. P. & Mutia, F. (2018). Library and Information (LIS) research topics in indonesia from 2006 to 2017. *Library Philosophy and Practice (e-journal)*. Retrieved from: <https://digitalcommons.unl.edu/libphilprac/1773>
- Baghmohammad, M., Mansouri, A. & Cheashmehsohrabi, M. (2020). Identification of topic development process of knowledge and information science field based on the topic modeling (LDA), *Iranian Journal of Information processing and Management*, 36(2), 297-328.
- Bai, X. (2018). Predicting the number of publications for scholarly networks. *IEEE Access*, 6, 11842–11848. <https://doi.org/10.1109/ACCESS.2018.2812804>
- Balili, C., Segev, A. & Lee, U. (2017, December). Tracking and predicting the evolution of research topics in scientific literature. In *2017 IEEE international conference on big data (big data)* (pp. 1694-1697). IEEE.
- Bengio, Y., Simard, P. & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2), 157-166. <https://doi.org/10.1109/72.279181>

- Bildosola, I., Gonzalez, P. & Moral, P. (2017). An approach for modelling and forecasting research activity related to an emerging technology. *Scientometrics*, 112(1), 557–572. <https://doi.org/10.1007/s11192-017-2381-3>
- Budi, A., Rizal, F. A. & Widodo, A. (2013). Prediction of Research Topics on Science & Technology (S&T) using Ensemble Forecasting. *International Journal of Software Engineering and Its Applications*, 7(5), 253-268. <https://doi.org/10.14257/ijseia.2013.7.5.23>
- Carlson, C. N. (2003). Information overload, retrieval strategies and internet user empowerment. Retrieved from http://eprints.rclis.org/5432/1/Information_Overload.pdf
- Chang, Y. W. & Huang, M. H. (2011). A Study of the Evolution of Interdisciplinarity in Library and Information Science: Using Three Bibliometric Methods. *Journal of the American Society For Information Science*, 63(1), 22-33. <https://doi.org/10.1002/asi.21649>
- Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3), 359–377. <https://doi.org/10.1002/asi.20317>
- Chen, C., Wang, Z., Li, W. & Sun, X. (2018, April). Modeling scientific influence for research trending topic prediction. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 32, No. 1).
- Cho, J. H., Lee, J. & Sohn, S. Y. (2021). Predicting future technological convergence patterns based on machine learning using link prediction. *Scientometrics*, 126(7), 5413–5429. <https://org/10.1007/s11192-021-03999-8>
- Choi, Y., Park, S. & Lee, S. (2021). Identifying emerging technologies to envision a future innovation ecosystem: A machine learning approach to patent data. *Scientometrics*, 126(7), 5431–5476. <https://doi.org/10.1007/s11192-021-04001-1>
- de Solla Price, D. (1965). Networks of scientific papers. *Science*, 149(3683), 510–515.
- Dora, M. & Kumar, H. A. (2019). National and international trends in library and information science research: A comparative review of the literature. *IFLA Journal*, 46(3), 234-249. <https://doi.org/10.1177%2F0340035219886610>
- Figuerola, C. G., Marco, F. J. G. & 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>
- Garner, J., Davidson, K. & Williams, V. K. (2008). Identifying serials trends through twenty years of NASIG conference proceedings: a content analysis. *Serials Review*, 34(2), 88-103. <https://doi.org/10.1080/00987913.2008.10765160>
- Ghanadinezhad, F. & Heidari, G. (2018). Identifying and analyzing research priorities in each of the research axes of knowledge and information science from the viewpoint of the faculty and doctoral students of this field. *Iranian Journal of Information processing and Management*, 34(1), 57-88. [in Persian]
- Godet, M. (2006). *Creating futures: Scenario planning as a strategic management tool*. Washington, DC: Economica. Economica Brookings diffusion.
- Guo, H., Weingart, S. & Borner, K. (2011). Mixed-indicators model for identifying emerging research areas. *Scientometrics*, 89(1), 421–435. <https://doi.org/10.1007/s11192-011-0433-7>

- Han, X. (2020). Evolution of research topics in LIS between 1996 and 2019: an analysis based on latent Dirichlet allocation topic model. *Scientometrics*, 125(3), 2561–2595. <https://doi.org/10.1007/s11192-020-03721-0>
- He, Q., Chen, B., Pei, J., Qiu, B., Mitra, P. & Giles, L. (2009, November). Detecting topic evolution in scientific literature: how can citations help?. In *Proceedings of the 18th ACM conference on Information and knowledge management* (pp. 957-966).
- Hessey, R. & Willett, P. (2012). Quantifying the value of knowledge exports from librarianship and information science research. *Journal of information science*, 39 (1), 141-150. <https://doi.org/10.1177%2F0165551512442476>
- Hochreiter, J. (1991). *Investigations on dynamic neural networks*. Diploma thesis, Technical University of Munich, Germany.
- Hoz-Correa, A., Munoz-Leiva, F. & Bakucz, M. (2018). Past themes and future trends in medical tourism research: A co-word analysis. *Tourism Management*, 65, 200-211. <https://doi.org/10.1016/j.tourman.2017.10.001>
- Joung, J. & Kim, K. (2017). Monitoring emerging technologies for technology planning using technical keyword based analysis from patent data. *Technological Forecasting and Social Change*, 114, 281–292. <https://doi.org/10.1016/j.techfore.2016.08.020>
- Kumar Sinha, M. (2016). Scenario of changing trends in library and information science education and research an analytical study of brochures of seminars, conferences and workshops organized during january 2012- july 2014. *Journal of Humanities and Social Science*, 21(7), 20-43.
- Lee, C., Kwon, O., Kim, M. & Kwon, D. (2018). Early identification of emerging technologies: A machine learning approach using multiple patent indicators. *Technological Forecasting and Social Change*, 127, 291–303. <https://doi.org/10.1016/j.techfore.2017.10.002>
- Li, X., Xie, Q., Daimb, T. & Huang, L. (2019). Forecasting technology trends using text mining of the gaps between science and technology: The case of perovskite solar cell technology. *Technological Forecasting & Social Change*, 146, 432-449. <https://doi.org/10.1016/j.techfore.2019.01.012>
- Liu, G. & Yang, L. (2019). Popular research topics in the recent journal publications of library and information science. *The Journal of Academic Librarianship*, 45(3), 278-287. <https://doi.org/10.1016/j.acalib.2019.04.001>
- Luo, L. & McKinney, M. (2015). JAL in the past decade: A comprehensive analysis of academic library research. *The Journal of Academic Librarianship*, 41(2), 123-129. <https://doi.org/10.1016/j.acalib.2015.01.003>
- Mansourkiaie, R., Babalhavaegi, F., Nooshinfard, F. & Soheili, F. (2019). Study of the Future of the Dissemination of Scientific Productions of Knowledge and Information Science in Social Networks from the Iranian Experts' Viewpoint. *Library and Information Sciences*, 22(3), 136-163. [in Persian]
- Mostafavi, I., Osareh, F. & Tavakolizadeh-Ravari, M. (2018). Identifying content structure of knowledge and information science (KIS) studies based on co-word analysis of articles in Web of Science (WoS) database (2009-2013). *Iranian Journal of Information processing and Management*, 33(3), 1271-1300. [in Persian]
- Olah, C. (2015). Understanding LSTM networks. Retrieved from <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

- Pestana, M. H., Sánchez, A. V. & Moutinho, L. (2019). The network science approach in determining the intellectual structure, emerging trends and future research opportunities – An application to senior tourism research. *Tourism Management Perspectives*, 31, 370-382. <https://doi.org/10.1016/j.tmp.2019.07.006>
- Porter, A. L., Garner, J., Carley, S. F. & Newman, N. C. (2018). Emergence scoring to identify frontier R&D topics and key players. *Technological Forecasting and Social Change*, 146, 628–643. <https://doi.org/10.1016/j.techfore.2018.04.016>
- Lara-Rodríguez, J. S., Rojas-Contreras, C. & Oliva, E. J. D. (2019). Discovering emerging research topics for brand personality: A bibliometric analysis. *Australasian Marketing Journal (AMJ)*, 27(4), 261-272. <https://doi.org/10.1016/j.ausmj.2019.06.002>
- Salatino, A. (2015). *Early Detection and Forecasting of Research Trends*. In *14th International Semantic Web Conference*, Bethlehem (PA), USA.
- Santo, M. M., Coelho, Santos, D. M. & Filho, L. F. (2006). Text mining as a valuable tool in foresight exercises: A study on nanotechnology. *Technological Forecasting & Social Change*, 73(8), 1013–1027. <http://dx.doi.org/10.1016%2Fj.techfore.2006.05.020>
- Sarah, M. & Pete, D. (2004). Striking a balance: Priorities for research in LIS. *Library Review*, 53(3), 167-176. <https://doi.org/10.1108/00242530410526574>
- Schwab, K. (2017). *The Fourth Industrial Revolution*. Switzerland: World Economic Forum.
- Singh, N., Hu, C. & Roehl, W. S. (2007). Text mining a decade of progress in hospitality human resource management research: Identifying emerging thematic development. *International Journal of Hospitality Management*, 26(1), 131-147. <https://doi.org/10.1016/j.ijhm.2005.10.002>
- Small, H. (2006). Tracking and predicting growth areas in science. *Scientometrics*, 68(3), 595-610. <https://doi.org/10.1007/s11192-006-0132-y>
- Small, H., Boyack, K. W. & Klavans, R. (2014). Identifying emerging topics in science and technology. *Research Policy*, 43(8), 1450-1467. <https://doi.org/10.1016/j.respol.2014.02.005>
- Soheili, F., Khasseh, A. & koranian, P. (2018). Thematic trends of concepts in Knowledge and Information Science based on co-word analysis in Iran. *National Studies on Librarianship and Information Organization (NASTINFO)*, 29(2), 171-190. [in Persian]
- Sohrabi, B. & Khalilijafarabad, A. (2018). Systematic method for finding emergence research areas as data quality. *Technological Forecasting and Social Change*, 137, 280-287. <https://doi.org/10.1016/j.techfore.2018.08.003>
- Summers, R., Oppenheim, CH., Meadows, J., McKnight, C. & Kinnell, M. (1999). Information Science in 2010: A Loughborough University View. *Journal of the American Society for Information Science*. 50(12), 1153-1162. [https://doi.org/10.1002/\(SICI\)1097-4571\(1999\)50:12%3C1153::AID-ASI20%3E3.0.CO;2-H](https://doi.org/10.1002/(SICI)1097-4571(1999)50:12%3C1153::AID-ASI20%3E3.0.CO;2-H)
- Tang, R. (2004). Evolution of the interdisciplinary characteristics of information and library science. *Proceedings of the American Society for Information Science and Technology*, 41 (1), 54-63. <https://doi.org/10.1002/meet.1450410107>
- Taşkın, Z. (2021). Forecasting the future of library and information science and its sub-fields. *Scientometrics*, 126, 1527–1551. <https://doi.org/10.1007/s11192-020-03800-2>
- Tavakoli, Gh., Babakirad, A. & Ghoroooneh, H. (2016). Development and designing of the method for emerging issues analysis. *Journal of Security Research*, 15(53), 151-181.

- Tuomaala, O., Järvelin, K. & Vakkari, P. (2014). Evolution of library and information science, 1965–2005: Content analysis of journal articles. *Journal of the Association for Information Science and Technology*, 65(7), 1446-1462. <https://doi.org/10.1002/asi.23034>
- Wu, F. S., Shiu, C. C., Lee, P. C., & Su, H. N. (2010, July). Integrated methodologies for mapping and forecasting science and technology trends: a case of etching technology. In *PICMET 2010 Technology Management For Global Economic Growth* (pp. 1-23). IEEE.
- Wu, Z.; Ren, Y. (2018). A bibliometric review of past trends and future prospects in urban heat island research from 1990 to 2017. *Environmental Reviews*, 27(2), 241-251. <https://doi.org/10.1139/er-2018-0029>
- Xiao, X., Zhang, F. & Li, J. (2015). Library and Information Science research in China: A survey based analysis of 10 lis educational institutes. *The Journal of Academic Librarianship*. 41(3), 330–340. <https://doi.org/10.1016/j.acalib.2015.02.012>
- Xu, S.; Hao, L.; An, X.; Pang, H.; Li, T. (2020). Review on emerging research topics with key-route main path analysis. *Scientometrics* 122, 607–624.
- Zhang, Y., Zhang, G., Chen, H., Porter, A. L., Zhu, D. & Lua, J. (2016). Topic analysis and forecasting for science, technology and innovation: Methodology with a case study focusing on big data research. *Technological Forecasting & Social Change*, 105, 179-191. <http://dx.doi.org/10.1016/j.techfore.2016.01.015>
- Zins, Ch. (2007). Knowledge map of information science. *Journal of the American Society for Information Science and Technology*. 58(4), 526-535. <https://doi.org/10.1002/asi.20505>
- Zong, Q. J., Shen, H. Z., Yuan, Q. J., Hu, X. W., Hou, Z. P. & Deng, S. G. (2013). Doctoral dissertations of Library and Information Science in China: A co-word analysis. *Scientometrics*, 94(2), 781-799. <https://doi.org/10.1007/s11192-012-0799-1>