

Mapping and Analyzing the Scientific Map of Knowledge Organization Using Research Indexed in the WOS Database

Ali Asghar Sadabadi*, Saeed Ramezani**, Kiarash Fartash*** and Iman Nikijoo****

*Shahid Beheshti University Tehran, Tehran 009821, Iran <a_sadabadi@sbu.ac.ir>

**Tehran University, Tehran, Iran <saeed.ramezani@ut.ac.ir>

*** Shahid Beheshti University Tehran, Tehran 009821, Iran <k_fartash@sbu.ac.ir>

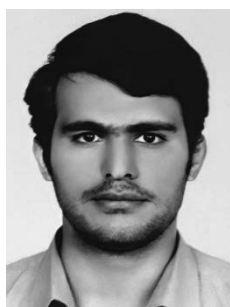
**** Shahid Beheshti University Tehran, Tehran 009821, Iran <i.nikijoo@mail.sbu.ac.ir>



Dr. Ali Asghar Sadabadi is currently the business incubator director of Shahid Beheshti University and assistant professor in the institute for science and technology studies at the SBU in Iran. He received his Ph.D. in science and technology policy from Tehran University in 2018. His areas of interests are social innovation, social entrepreneurship, knowledge management and business management, and he has written more than 50 research articles in these fields.



Dr. Saeed Ramezani received his Ph.D. degree (Tehran University) in Systems Management. He is a lecturer and researcher in the field of industrial engineering, management systems and technology management. He has published over 40 books and research articles. As a lead auditor he has been approved by some international companies in ISO and Products standards. He executed more than 60 major management projects in Iran, Turkey, UAE.



Kiarash Fartash holds a BSc in Electrical Engineering, an MBA and a PhD degree in Management of Technology with specialization in S&T Policy from Allameh Tabataba'i University. He is an assistant professor at Shahid Beheshti University in Tehran and has authored more than 30 research papers. He has more than a decade of experience in the capacity of STI policy researcher as well as policy officer. Kiarash's research interest entails innovation policy, energy policy and sustainability transition.



Iman Nikijoo received his M.Sc. degree in technology management from the faculty of management and accounting at Shahid Beheshti University, Iran in 2022. His current research interests are social innovation, social entrepreneurship, sustainability transition social network analysis.

Sadabadi, Ali Asghar, Saeed Ramezani, Kiarash Fartash and Iman Nikijoo. 2022. "Mapping and Analyzing the Scientific Map of Knowledge Organization Using Research Indexed in the WOS Database." *Knowledge Organization* 49(6): 448-464. 63 references. DOI:10.5771/0943-7444-2022-6-448.

Abstract: Scientometrics has found many applications in describing, explaining and predicting the scientific status of researchers, educational and research groups, universities, organizations and countries in various national and international arenas. By studying the scientific products of different countries, their status in the production of science can be evaluated. Present study was conducted using a scientometrics approach and using co-word analysis and social network analysis (SNA) to investigate relationships in the field of knowledge organization. In this regard, research indexed in web of science on the topic of "knowledge organization" has been analyzed using software including VOSviewer, Gephi, Publish or Perish. The findings of the study show that the most frequently used topics and words are knowledge organization and classification. Also, the most valuable subject areas were identified based on the maps drawn using the closeness and centrality of indexes, taxonomy, ontology and knowledge organization systems. Co-authorship analysis revealed that the co-authorship network is discrete and has low-density, with a total of 12,491 citations in all articles. Also, the most prolific author is Hjørland, followed by Smiraglia and Dahlberg. using the co-word map of knowledge organization, policymakers can plan appropriately through the knowledge of the research and thematic status of knowledge organization.

Received 8 November 2020; Revised 1 July 2022, 19 August 2022; Accepted 23 August 2022

Keywords: knowledge organization, scientific map, co-citation analysis, co-authorship, social network analysis

1.0 Introduction

Today, one of the most validated methods for evaluating scientific output is the use of scientometric indices, commonly referred to as the knowledge of measuring science, which has emerged in the scientific community since the late 1960s (Bookstein 1994).

Knowledge organization is about providing optimal conditions for the identification and retrieval of documents or parts of documents (Broughton et al. 2005). Guimarães defines knowledge organization, the study and development of the fundamentals and techniques of planning, construction, management, use and evaluation of systems of description, cataloguing, ordering, classification, storage, communication and retrieval of documents created by man to witness, conserve and transmit their knowledge and actions, based on their content, in order to ensure their conversion into information capable of generating new knowledge. He believed knowledge organization to be a three-dimensional science, as it deals with the principles, methods and instruments put into action for the management of human knowledge from a triple perspective: its representation, its organization and its documental communication (Guimarães 2014).

Castanha and Wolfram (2018) believe the domain of knowledge organization is in continuous development and it is also a relatively autonomous, interdisciplinary space that interfaces with other subject areas. They state that knowledge organization is concerned with issues of a theoretical and methodological nature that contribute to the systematization, production, organization, dissemination, representation and retrieval of information in different scholarly contexts.

Within information science, knowledge organization stands as a mediator core between the production and use of information (Guimarães et al. 2008) and configures a set of processes that establish the mediation between knowledge, that once produced is embodied and socialized in such a manner that it can serve as a foundation for the following generation of new knowledge. What can be called a spiral flow of information (Castanha and Wolfram 2018).

For Dahlberg (1993), the information needs to be organized to become knowledge. Also, the knowledge organization should be based on knowledge units: established concepts that do not work if isolated but do if related (Gouvêa Meireles et al. 2014).

Knowledge organization, at least as it is practised within the domain of library and information science, has been largely (up to now) the province of the construction of tools for the storage and retrieval of documentary entities. That is, tools, such as catalogs, indexes, and databases, have been constructed to allow the rapid manipulation of and retrieval from large collections of surrogate records that represent

documents, which in turn represent recorded knowledge (Smiraglia 2002).

Given the increasing volume and substantial production of scientific information in various specialized fields, it is difficult to review all scientific literature produced by experts. Therefore, the use of international indexes and databases containing a significant portion of internationally valid information can be the fastest way to access information in any field (Hamidi et al. 2008). In this regard, Web of Science Database research on the topic of “knowledge organization” has been analyzed from the beginning using scientometrics and social network analysis software programs including VOSviewer, Gephi, Publish or Perish.

Due to the importance of the subject, we decided to pursue the following using scientometrics in this field:

1. A deeper analysis of this area
2. Extracting keywords and sub-categories of this field
3. Co-authorship
4. Provide value to stakeholders

2.0 Research background

With the advent and variety of scientometrics techniques, it has become possible for researchers to explore the intellectual structure of their preferred disciplines depending on the method used (Lane 2010). In 1986, Callon et al. published a book entitled *Mapping the Dynamics of Science and Technology: Sociology of Science in the Real World*, which is a prominent work in co-word analysis. Many researchers have used this method to explore conceptual networks in various fields of science.

As a systematized re-creation of a research process, bibliometric studies can reveal snapshots of the perception of a domain's research patterns (Martínez-Ávila et al. 2015). Knowledge organization (KO) is about describing, representing, filing and organizing documents and document representations as well as subjects and concepts both by humans and by computer programs.

KO has been defined as “the domain in which the order of knowledge is both the primary paradigm for scientific investigation and the primary application in the development of systems (Smiraglia 2002).

KO is a field of research, teaching and practice, which is mostly affiliated with library and information science (LIS), with KO understood in narrow senses, as well as in broad senses. The narrow senses are, for example, the Knowledge Organization Systems (KOSs) and Knowledge Organization Processes (KOPs) which occur within LIS. The broad senses are, for example, the conceptual systems, the social fields, and the activity systems, existing or taking place, in all spheres of society (Hjørland 2016).

In KO the concept of a knowledge organization system (KOS) is a generic term used for authority lists, classification systems, thesauri, topic maps, ontologies, etc. (Hodge 2000). A KOS can be defined as a selected set of concepts together with an indication of (some of) their semantic relations (Hjørland 2013).

For these purposes, rules and standards are developed, including classification systems, lists of subject headings, thesauri and other forms of metadata. The organization of knowledge into classification systems and concept systems are core subjects in KO. The two main aspects of KO are (1) knowledge organization processes (KOPs) and (2) knowledge organization systems (KOSs). Knowledge organization processes are, for example, the processes of cataloging, subject analysis, indexing, tagging and classification by humans or computers. Knowledge organization systems are the selection of concepts with an indication of selected semantic relations. Examples are classification systems, lists of subject headings, thesauri, ontologies and other systems of metadata.

The development of KO as a discipline for research and teaching is mainly tied to the development of library and information science as a university discipline (or professional school discipline), that is after 1850.

Some of the scientometrics research and knowledge organization using word co-occurrence, co-authorship, and co-citation are listed in Table 1.

Knowledge organization and bibliometrics have traditionally been seen as separate subfields of library and information science, but bibliometric techniques make it possible to identify candidate terms for thesauri, and to organize knowledge by relating scientific papers and authors to each other and thereby indicate kinds of relatedness and semantic distance. It is therefore important to view bibliometric techniques as a family of approaches to KO in order to illustrate their relative strengths and weaknesses. Using citation-based methods as a complement or alternative to conventional approaches to KO is thus not new in the bibliometric community, but rather tends to be neglected by the KO community (Hjørland 2013).

Whereas research in knowledge organization, as well as in all other scientific areas, results in a dynamic context of production and elaboration of knowledge, this context is continuously transformed because of its practical application and scientific communication (Castanha and Wolfram 2018).

There seems to be a need for more research on the co-word analysis method in knowledge organization. Therefore, this research aims to identify and map the knowledge organization research field through co-word, co-citation and co-authorship methods and to determine the efficiency of this method in identifying and defining the scientific and research priorities of this field.

3.0 Methodology

The fields of KO and bibliometrics have so far not had much mutual contact but today bibliometric mapping is a dominant method in the study of relations between research fields. Bibliometrics (with altmetrics, informetrics, scientometrics and webometrics) is an interdisciplinary field with strong affiliations to LIS. This field developed techniques for producing bibliometric maps based on co-citation analysis, bibliographic coupling, or by direct citation. Such maps may serve information retrieval and are a form of alternative or a supplementary approach to knowledge organization (Hjørland 2016).

In this context, an analysis and assessment of scientific production has become essential to creating instruments for identifying behavior within a scientific field. This highlights the role of bibliometrics in making a significant contribution by providing quantitative analyses of the communication processes, and of the nature and development of scientific domains that allow an objective and broad view of a scientific domain. Among the methodologies for evaluation of science, bibliometrics stands out.

Scientometrics is one of the most important scales for evaluating scientific products. Macías-Chapula (2004) argues that “scientometric indices have become essential for the estimation of the modern state (in line with the latest developments) of a particular subject for the scientific community.” Scientometrics deals with, and overlaps with, bibliometrics and information science interests. The terms bibliometrics, scientometrics, and information science refer to the component parts of studying the dynamics of the fields as reflected in the production of their sources (Mooghali 2011). Scientometrics can be defined as “the quantitative study of science, communication in science, and scientific policy.” What began as Eugene Garfield's idea of creating an index to improve information retrieval in the 1960s and led to the creation of the Science Citation Index (SCI) soon became a new tool in the empirical study of science (Leydesdorff 2015).

Van Ran (1997) believes that scientometrics research is limited to quantitative studies of science and technology. It aims at advancing knowledge, technology and science; it is also about social and political issues. He divides the main interests of scientometrics research into four intertwined areas:

1. development of methods and techniques for designing, making and using quantitative indicators in important aspects of science and technology;
2. development of information systems in science and technology;
3. the study of the interaction between science and technology;

ID	Article title	Authors	Year	
	Scientometrics and patent bibliometrics in RUL analysis	V. Cavaller	2009	The analysis of the life cycle allows the incorporation of qualitative considerations (legal, contractual, physical, technical know-how, functional, economic) related to the conduct of future technologies.
	Global trend of open innovation research: A bibliometric analysis	Thu Le et al.	2019	A bibliometric approach was conducted to evaluate the global scientific outputs of open innovation based on literature in Social Science Citation Index (SSCI) database from 2003 to 2017.
	Major trends in knowledge management research: a bibliometric study	Akhavan et al.	2016	This study provides an overview of the knowledge management literature from 1980 through 2014. It employs bibliometric and text mining analyzes on a sample of 500 most cited articles to examine the impact of factors such as number of authors, references, pages, and keywords on the number of citations that they received.
	A new method for automatically constructing domain-oriented term taxonomy based on weighted word co-occurrence analysis	Li et al.	2015	It introduces the definition of weighted co-occurring word pair and corresponding improved method of word co-occurrence analysis.
	An application of co-word analysis and bibliometric maps for detecting the most highlighting themes in the consumer behavior research from a longitudinal perspective	Muñoz-Leiva et al.	2011	This article presents an application of a bibliometric and visual study of the research carried out on a social science subfield, concretely the consumer behavior research (CBR), from a longitudinal perspective (period 1966–2008).
	The domain of knowledge organization: A bibliometric analysis of prolific authors and their intellectual space	Castanha and Wolfram	2018	Considering the importance of identifying and studying the science dynamics of KO in the context of information science, this study aims to analyze the most productive authors contributing to the journal <i>Knowledge Organization</i> over the period 1993-2016, along with the authors' references and citations received. More specifically, this study performs a citation, co-citation a bibliographic coupling analysis of the KO journal literature in order to verify the dynamics of the KO area using these indicators. Also, researchers analyze the theoretical references and most useful themes that have constituted this scholarly community. on this basis, Birger Hjørland was the most cited author, and was situated at or near the middle of each of the maps based on different citation relationships.
	The intellectual landscape of the domain of culture and ethics in knowledge organization: an analysis of influential authors and works	Wang	2019	This article aims to examine the intellectual landscape of the domain of culture and ethics in knowledge organization (KO). A domain analysis was conducted on a corpus of 206 relevant papers using bibliometric methods such as author co-citation analysis. The findings revealed a core group of influential authors consisting of Olson, Beghtol, and other influential KO researchers. The most prominent research themes that constitute this scientific community have also been investigated by reviewing some fundamental concepts and influential works in the domain.
	The Representation of Ethics and Knowledge Organization in the WoS and LISTA Databases	Martínez-Ávila et al	2015	This paper describes and analyzes the results of our bibliometric analysis studying the representation of ethics and KO in WOS and Library, Information Science & Technology Abstracts (LISTA) databases. The motivation of the study is to re-create how an average or naive user might perceive the topic of knowledge organization and ethics through the search results in these two common research tools. authors aim to highlight how the omissions and insufficiencies of these databases provide an incomplete picture when superficially researching these topics
	Evolution of research activities and intellectual influences in information science 1996–2005: Introducing author bibliographic-coupling analysis	Zhao and Strotmann	2008	Researchers updated the White and McCain (1998) study on information science for the years 1996–2005 (Hjørland, 2013). They found that Author Bibliographic coupling analysis is an effective method for providing a realistic picture of current active research within a research field, whereas Author co-citation analysis studies the external and internal as well as recent and historical intellectual influences on the field. Also, they stated When combining ACA with ABCA, it is possible to gain a wide view of the intellectual structure of a research field.

Table 1. Literature review in the field of scientometrics and knowledge organization.

- investigating the cognitive and socio-organizational structures of scientific domains and development processes in relation to social factors (Mooghali 2011).

Understanding the impact of scientometrics on the development of academic disciplines is a complex issue that is of great importance. Its relevance is due to the recent trend to

introduce numerical scales of scientific performance to evaluate research activities and facilitate comparison at different levels, for example, among researchers, institutions, or the media. These comparisons usually take the form of ratings designed to “measure” the “quality” of universities, scientists, scientific articles, and journals on a single scale. This development also points to the methodological shift within scientometrics, whose conceptual origins lie in an interpretive analysis of scientific communication aimed at understanding the characteristics of the discourse of academic publishing (Aistleitner 2018).

Social network analysis researchers understand networks as sets of objects called “nodes” that are connected by one or more relationships called “edge”. In the social sciences context, nodes can be a wide range of social units such as individuals, groups, organizations, corporations, governments, and countries (Buch-Hansen 2013). This applied research has been conducted using common methods in the study of scientometrics with techniques of co-word analysis, co-citation analysis, co-authorship analysis and social network analysis. Articles in the field of knowledge organization indexed in the Web of Science database from 1979 to 2022, are analyzed in this research. For this reason, the word “knowledge organization” was searched in the titles, abstracts and keywords articles.

After retrieving records related to knowledge organization, for co-word analysis, 1166 records were recovered. In the next step, common words were examined using expert opinions, and some words that had no domain-specific meaning in the structure of their respective sciences, such as history, model, quality and etc. were removed along with country names like Brazil.

Next, keywords with a frequency of 5 and more were included in the final analysis. It should be noted that different studies using co-word analysis have used different thresholds for inclusion of top keywords in the final analysis. For example, Liu et al. (2012) limited their analysis to 66 keywords that accounted for about 55% of the total frequency (Liu et al. 2012). Other features of the co-word matrix network such as centrality, density, weight, etc. were measured using Gephi software to find out more about the study area. In addition, in co-authorship analysis among the 2095 unique authors in knowledge organization research, it was identified that 27 authors had participated in at least 5 studies. It was further observed in the co-citation analysis that a total of 35,665 citations were made in all articles.

4.0 Findings

The results of 1166 articles obtained from the Web of Science, which have been analyzed in the Publish or Perish software, are as follows:

Metrics

Reference date:	2022-06-19 14:38:30 +0430
Publication years:	1979-2022
Citation years:	43 (1979-2022)
Papers:	1166
Citations:	12491
Citations/year:	290.49 (acc1=349, acc2=178, acc5=41, acc10=8, acc20=5)
Citations/paper:	10.71

According to the data, the total number of articles reviewed on the subject of knowledge organization has been 1166 studies since the beginning, indexed in the web of science database over the past 43 years, during which 12491 were cited, i.e., 10.71 times per article and 290.49 citations per year, which indicates the high quality of the research conducted.

4.1 Research trend in the field of intangible assets

The results of the analysis of the studies show that the research and development of scientific products in the field of knowledge organization in the web of science database has been very positive since the beginning, such that in the last 10 years the growth rate of scientific products is 75% of all research carried out from 1979 to 2022.

Among the most important channels of information exchange in scientific disciplines are the scientific journals of each discipline. Therefore, identifying active and reputable journals in any field is of particular importance. As shown in Figure 2, a review of data related to the contribution of each journal to the published articles indicates that “Nomos Verlagsgesellschaft” has the largest number of articles and research with 124 articles.

Figure 3 shows that about 77% of the studies were articles.

4.2 Word analysis

4.2.1 Keywords in the field of knowledge organization

As noted, for the purpose of co-word analysis, by placing inclusion thresholds on keywords that have been repeated at least 5 times, 177 frequent keywords were identified, of which, after eliminating irrelevant items, 92 remained to be included in the final analysis. The keywords are shown in Table 2 along with the degree and weights obtained. “Classification” has the highest number of connections, indicating the closeness to the knowledge organization concept. After “classification”, “knowledge” and “ontology” have gained the highest weight in the word co-occurrence network.

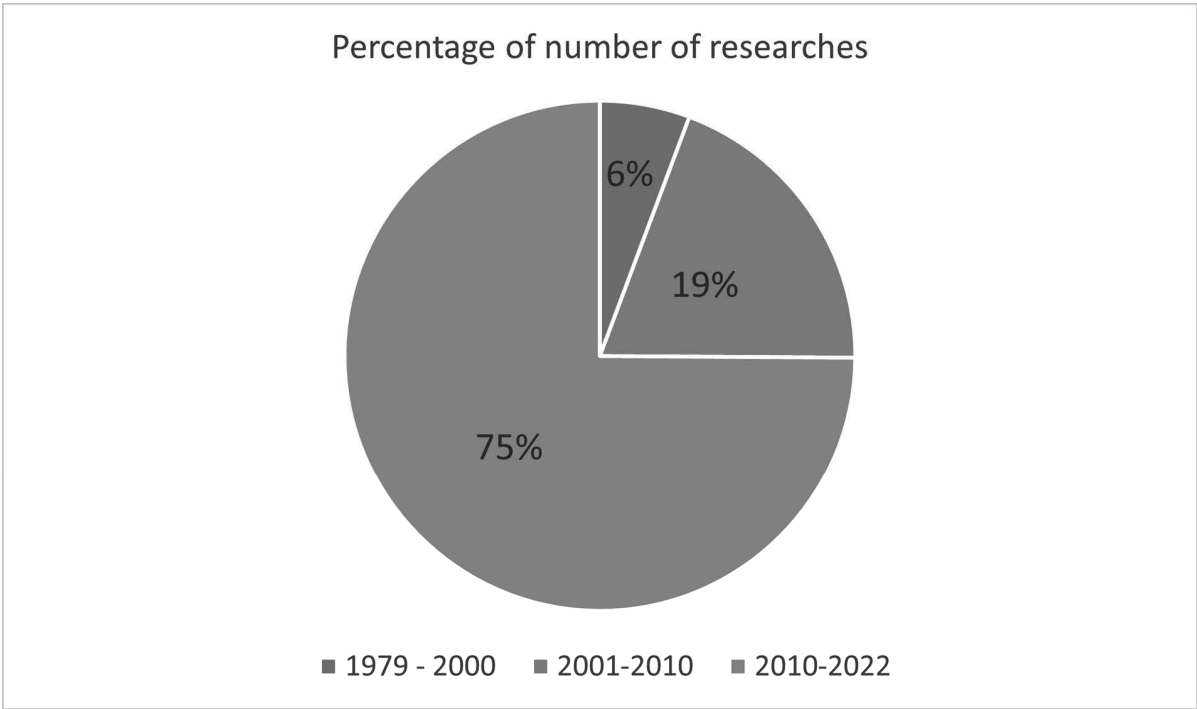


Figure 1. Papers from 1979 to 2022.

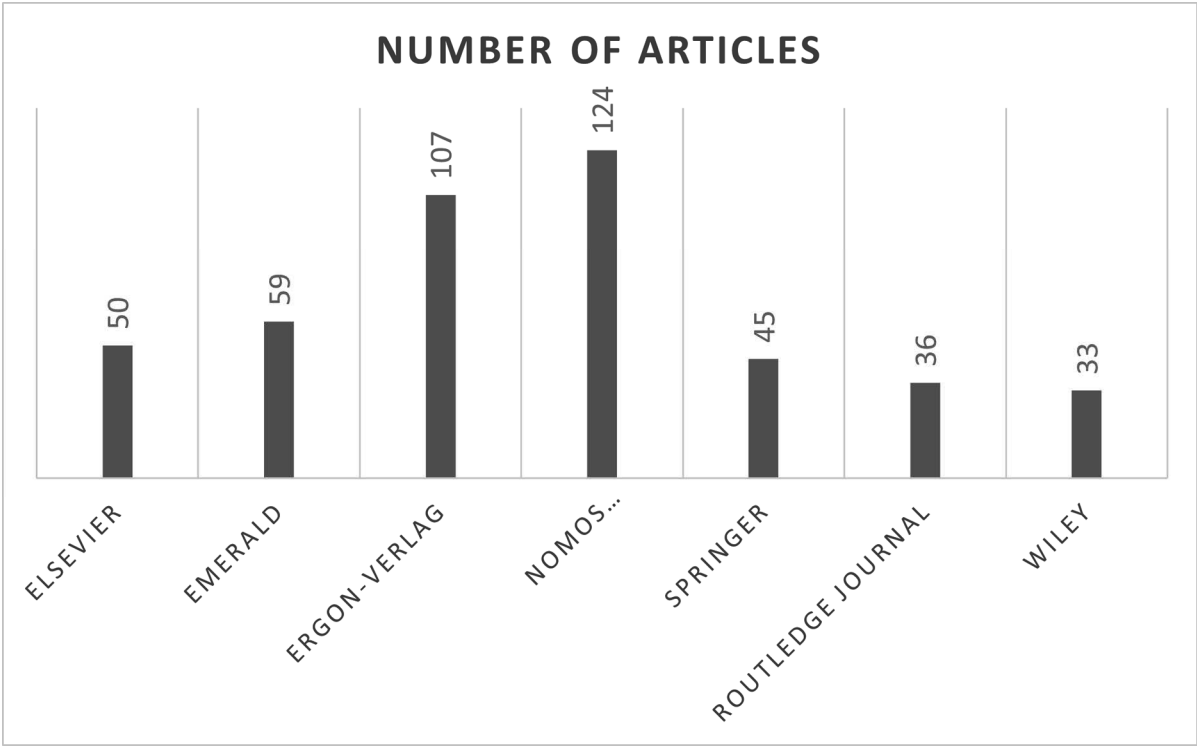


Figure 2. Contribution of each publisher to research.

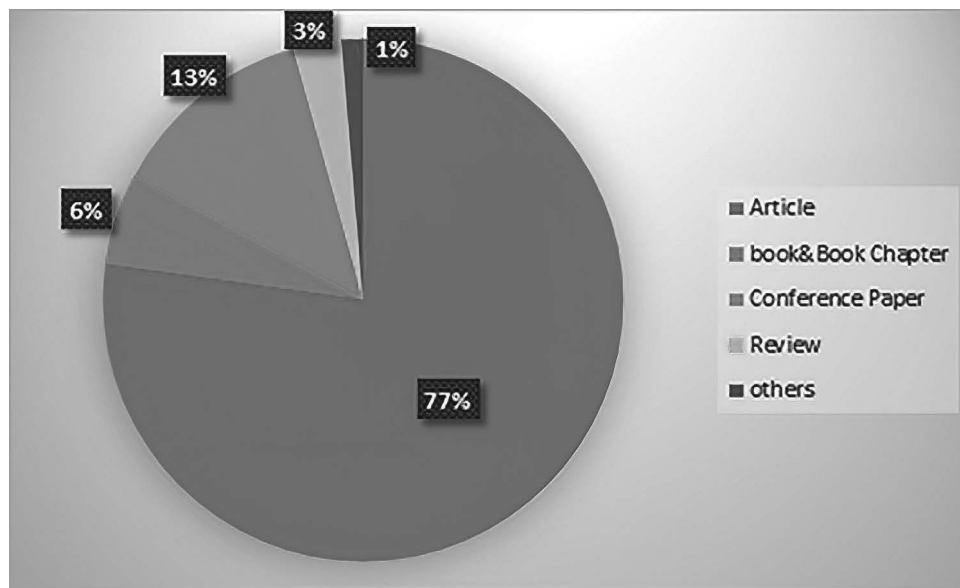


Figure 3. The percentage of the types of research.

Label	In degree	Out degree	Degree	Label	Weighted in degree	Weighted out degree	Weighted degree
classification	4	68	72	classification	10	332	342
knowledge	32	28	60	knowledge	91	107	198
ontology	33	18	51	library	129	24	153
library	33	16	49	ontology	95	32	127
knowledge organization systems	25	22	47	information science	59	66	125
semantic web	31	7	38	knowledge organization systems	60	45	105
information science	8	29	37	semantic web	73	22	95
framework	19	18	37	domain analysis	9	82	91
epistemology	8	27	35	epistemology	19	56	75

Table 2. Keywords co-occurrence network analysis.

4.2.2 The most frequent keywords

Words such as classification, knowledge, library, ontology, and information science were highly frequent in the present study (Table 3).

4.2.3 The most related words

The important “centrality or degree criterion” index is related to the number of edges that each node in a network is associated with. Therefore, the degree criterion deals with the role of each node in the network. In networks with directed edges, in-degree is the number of edges a node re-

ceives, while the out-degree is the number of edges that each node sends out. The higher the node's in-degree (especially, compared to the nodes that have a high level of in-degree), the stronger or more important the node will be (Scott 2011). Table 4 shows the keywords with the highest degree and, as is evident, classification and knowledge hold highest degrees.

4.2.4 Mapping co-occurrence

We used social network analysis techniques to interpret scientific maps. Because scientific maps have a structure similar to social networks (Guns et al. 2011). After modifying the network and eliminating unrelated nodes such as coun-

Keyword	Occurrences	Total link strength
classification	137	564
knowledge	100	362
library	53	235
ontology	57	197
information science	53	191

Table 3. The most frequent words in the co-occurrence network.

Label	Degree	Weighted degree
classification	72	342
knowledge	60	198
ontology	51	127
library	49	153
knowledge organization systems	47	105
semantic web	38	95
information science	37	125

Table 4. Words with high degree centrality.

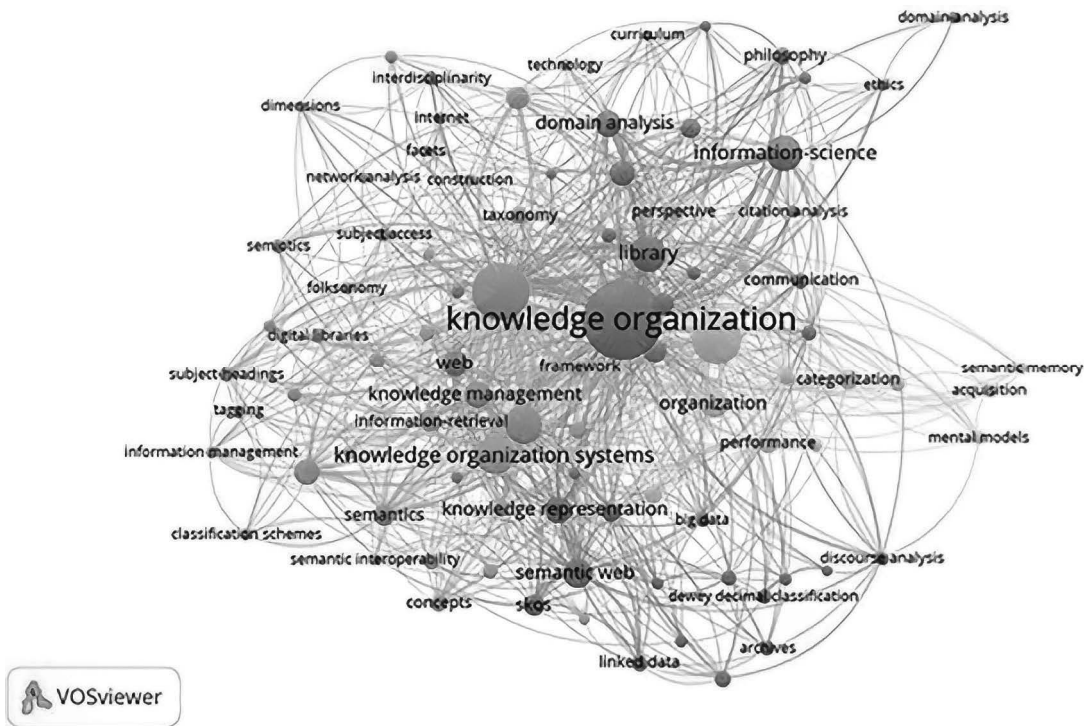


Figure 4. Word co-occurrence map.

tries, methodology related concepts, and so on, the final co-occurrence map was drawn. The map of concepts related to knowledge organization arising from the keywords of the research are drawn in the general form of word density using

VOSviewer software (Figure 4) and degree rank (Figure 5). VOSviewer software assigns each keyword a density based on frequency, and words with higher frequencies based on co-word analysis will have more densities. The purpose of

Gephi software is also to study the correlation of node properties and network construction using visual patterns. Social network analysis metrics such as node degree or betweenness can be used for better computation and visualization.

The longest path in the network, or network diameter, is 4 and the average distance of two nodes is calculated 1.79. The diameter of the network is measured by the distance of the longest paths in the network with the distance of the shortest paths (in terms of number of edges or connections) between two nodes. The shorter the network diameter, the faster the communication speed will be (Cheong and Corbitt 2009), and in the present network each two words can be connected through four intermediaries.

4.2.5 Co-occurrence network analysis

There are various indicators for word co-occurrence network analysis that can be applied to scientific maps. For example, the size of the network is determined by the number of nodes and the density of the network by the number of edges in the network. Centrality is one of the important indicators in network analysis. This index refers to the position of specific nodes within the network and its types are closeness, betweenness and degree centrality. Degree centrality is the simplest type of centrality is the value of each node obtained by counting the number of its neighbors. The number of neighbors is obtained based on the edges to which the node is connected. In a word co-occurrence network, the more centrality degree a word has, the more communication and network it has and the more effective it will be. Closeness centrality is based on the concept of distance and path length. In a network, vertices that have the least distance to all other vertices have a higher closeness centrality (Degenne and Forsé 1999). In the present study, closeness centrality is the words that link with the least distance (least number of edges) to other words. betweenness centrality also indicates the importance of the node in terms of its location on the map and in terms of information transmission over the network. Map analysis shows that there are 92 nodes and 907 edges in this map. Because the number of edges is more than nodes, the network is therefore continuous.

4.2.6 Centrality

The centrality of the network nodes can be studied by using the three indices of degree, betweenness and closeness. The degree centrality for each node in the network is equal to the sum of the edges entering the node, and closeness is the average length of the shortest paths between that node and the other nodes in the network. Betweenness refers to the number of times that a node is placed on the shortest paths between each two nodes in the network (Opsahl et al. 2010).

4.3 Degree centrality

Centrality is one of the network metrics or indicators that is useful in analyzing the structure of entire networks and node positions in the network. It measures the number of edges going in or out of a node in a network. The degree centrality is the number of edges a node shares with other nodes in the network. A subject with the most edges has the highest degree and is the most centralized node (Bródka et al. 2011). Degree centrality can facilitate or prevent the flow of resources between nodes in the network (Estrada and Rodriguez-Velazquez 2005). As shown in Figure 4, two nodes or keywords are connected if they share at least one common co-occurrence in this network. The size of each node represents the degree centrality index or the number of co-occurrences of that node with the other nodes in the network. Moreover, the more the co-occurrence of two nodes, the greater the diameter of the link between the two nodes and the bolder the lines that connect them. In the present network, after the keyword of the knowledge organization, the keyword classification has the highest degree centrality, which actually represents the second highest number of connections with other nodes.

4.3.1 Betweenness centrality

Betweenness centrality as a node compatibility feature indicates the importance of the node in terms of its position on the map and in terms of information on the network. The betweenness centrality index is calculated based on the position of people in the network. The node that has the highest betweenness centrality is in the middle of a large number of other nodes and the edges connecting other nodes pass through it. These nodes have the power to isolate and enhance communication (Newman 2005). Analysis of betweenness shows that most keywords have a betweenness of less than 100, except for 6 nodes.

Based on betweenness analysis, results shown in table 5 are the most important subjects in information transition in the network. In other words, these subjects are considered to be interdisciplinary in relation to the field of knowledge organization. Table 5 shows keywords with highest betweenness centrality.

4.3.2 Closeness centrality

How fast an entity on the network can access more entities on that network is explained by closeness. An entity with higher closeness generally has the following characteristics:

It has quick access to other entities on the network; it has a short path to other entities; it is close to other entities; and it has a high visibility about what is happening on the network (Visualizer 2009). Closeness is calculated based on the

Label	Betweenness centrality
knowledge	208.49
classification	168.78
ontology	149.30
library	133.27
framework	96.22
information science	78.71
epistemology	76.58
information retrieval	68.37

Table 5. Words with highest betweenness centrality.

Label	Closeness centrality
taxonomy	1
SKOS	1
subject access	1
standards	1
library and information science	1
ontology	0.95
semantics	0.916667
classification	0.909639

Table 6. Keywords with highest closeness centrality.

geodesic distance. It measures the distance a node has from other nodes. This index represents the availability (Frank 2002). According to the network, most nodes are close to one another and have a closeness between 0 and 1, which allows information to spread very quickly across the network.

According to calculations, topics like taxonomy, SKOS, subject access, and standards have the highest impact and the closeness centrality. In fact, such nodes are less distant from all other nodes and are, on average, closer to all nodes, and in fact these nodes are more robust in the network (Crucitti et al. 2006; Okamoto et al. 2008). Table 6 shows the keywords with the highest closeness centrality:

4.3.3 Co-occurrence network density

Density is one of the indicators used to check the degree of network cohesion. Network density can be defined as a set of edges that connect nodes to each other and prevent the network from breaking (Faust 2006). Regarding network

cohesion, it can be stated that if the interconnection between the nodes or the number of edges between the nodes is low, the network has low cohesion and there will be many holes in the network. In this case, the network will be discrete and information flow will be very slow. Conversely, when there is a high number of edges between the nodes and the holes in the network are few, the network is continuous (Faust 2006; Kohler et al. 2001). Analysis of word co-occurrence shows that the network has relatively low cohesion due to its density, which is 0.108. This density indicates that only 10.8% of the potential internal communications in the network is realized. In other words, the number of edges available in the network can be relatively low, resulting in a slow information flow.

The network density also determines the ratio of the number of connections available to the number of connections possible in the network. The co-word density map shows in Figure 5.

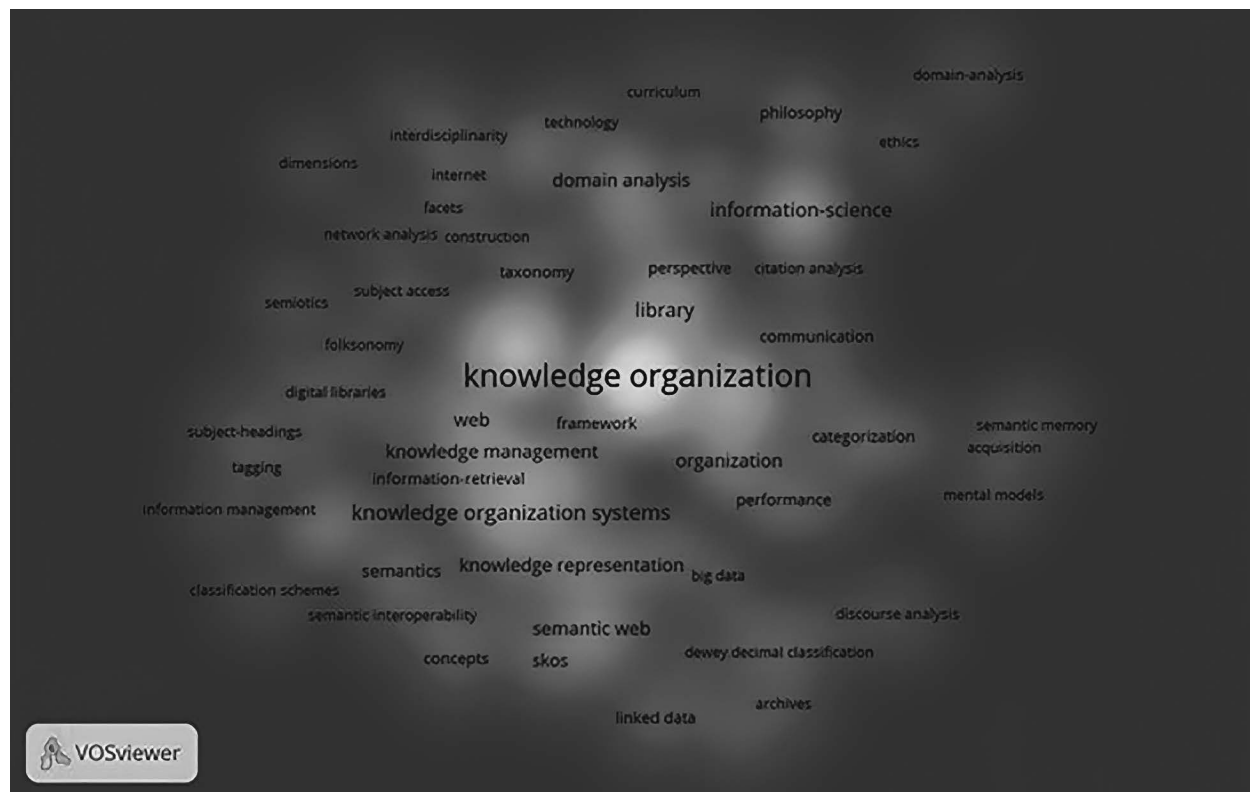


Figure 5. Density based word co-occurrence network.

4.3.4 Hub

Network elements can be categorized into different classes through their participation in the network. For example, they are classified according to the way they communicate with other components of the system. One of the important distinctions that can be made based on their effect is the potential effect they have on the overall system and their capacity to transmit or process information. High-impact nodes are often referred to as hubs. Identification of hubs is needed to map the different areas that are most interoperable and contribute to system integration. One issue to note is that nodes that have high connections also have a greater shareability factor. The high degree of centrality and their relevance as well as their high degree of interconnectedness in structural communication indicate that they play a vital role in the integration of processes and information flows (Franks et al. 2008).

Nodes that act as hubs and connect many nodes. Table 7 shows that nodes of classification, knowledge and framework are the largest network hubs.

4.3.5 Cluster analysis

Clustering means categorizing objects in collections without supervision and interference. In this way, clusters are

not predetermined and cluster labels are not available. Clustering refers to finding a structure within a set of unlabeled data; it is said to be a collection of data-points that are similar to one another (Jain et al. 1999). Classification and clustering are different. In a classification, you assign an object a predefined class; but in a cluster, there is no knowledge of the classes within the data, and the data itself is the source of clusters (Omran et al. 2007). Consequently, after performing clustering, an expert should interpret the clusters, and, in some cases, it may be necessary to reconsider some of the parameters after clustering. These are the parameters included in the clustering task but are irrelevant or of little importance and are therefore eliminated and clustering is performed again (Edelstein 1999).

Cluster analysis was performed using Gephi and 3 clusters of topics were identified and clustered by experts. The “knowledge organization” cluster is the largest cluster with 47 points and contains clusters related to classification and knowledge.

The clustering results in Table 8 show the number of clusters, their naming and the percentage of concepts present in each cluster. The clustering shows the automatic structuring of the concepts, and the concepts within each group are the most similar, and all 3 clusters have the least intra-group differences and the maximum inter-group differences. In the “knowledge organization” cluster, for exam-

Label	Hub
classification	0.359352
knowledge	0.230165
framework	0.217609
design	0.208714
epistemology	0.197774
information science	0.18846
domain analysis	0.185167
knowledge organization systems	0.172003
information retrieval	0.159092

Table 7. Main hubs of word co-occurrence network.

Network attributes				
Average path length	Nodes	Average clustering coefficient	Edge	
1.794	92	0.284	907	
Share of network		Number of nodes	Cluster	#
51%		47	knowledge organization	0
25%		23	knowledge organization systems	1
23%		22	ontology	2

Table 8. Network attributes.

ple, the concepts of classification, knowledge, categorization, and library are like most closely related.

The clustering coefficient shows how much the keywords tend to create co-occurring clusters. The clustering coefficient is always a number between zero and one (Miguel et al. 2012). A mean of 0.284 for clustering coefficients for keyword clustering was obtained which indicates that the vocabulary is not randomly formed in a cohesive network, and the coefficient estimates the tendency of network nodes to communicate with each other and form clusters. Moreover, the average distance of the network nodes was 1.794, which indicates the closeness of the nodes and the density of the network.

4.3.6 Authors and co-authors (co-authorship)

There are 2095 unique authors in the knowledge organization area. This is obtained from all of the knowledge organization research since the beginning which is indexed on the web of science database (Figure 6).

The co-authorship network of authors contributing at least 3 works has 28 nodes and 44 edges, indicating a low co-authorship.

By computing the co-authorship network density of knowledge organization network, the value obtained was 0.058. The network density shows the number of edges available in the network in relation to the number of possible edges. Network density is a number between 0 and 1. The closer a number is to 1, the more inter-node connections there are in a network. The higher the density of a network, the more nodes are connected and the closer they are (Mrvar and Batagelj 2016). Overall, the obtained number indicates that the co-authorship network is discrete and has a low density.

There were 355 single author studies, 278 two-author studies, 160 three-author studies, and the rest of studies have higher author numbers. Careful analysis of the number of authors per article indicates that the dominant authoring pattern in knowledge organization research is collective, and only about 35% of the articles have a single au-

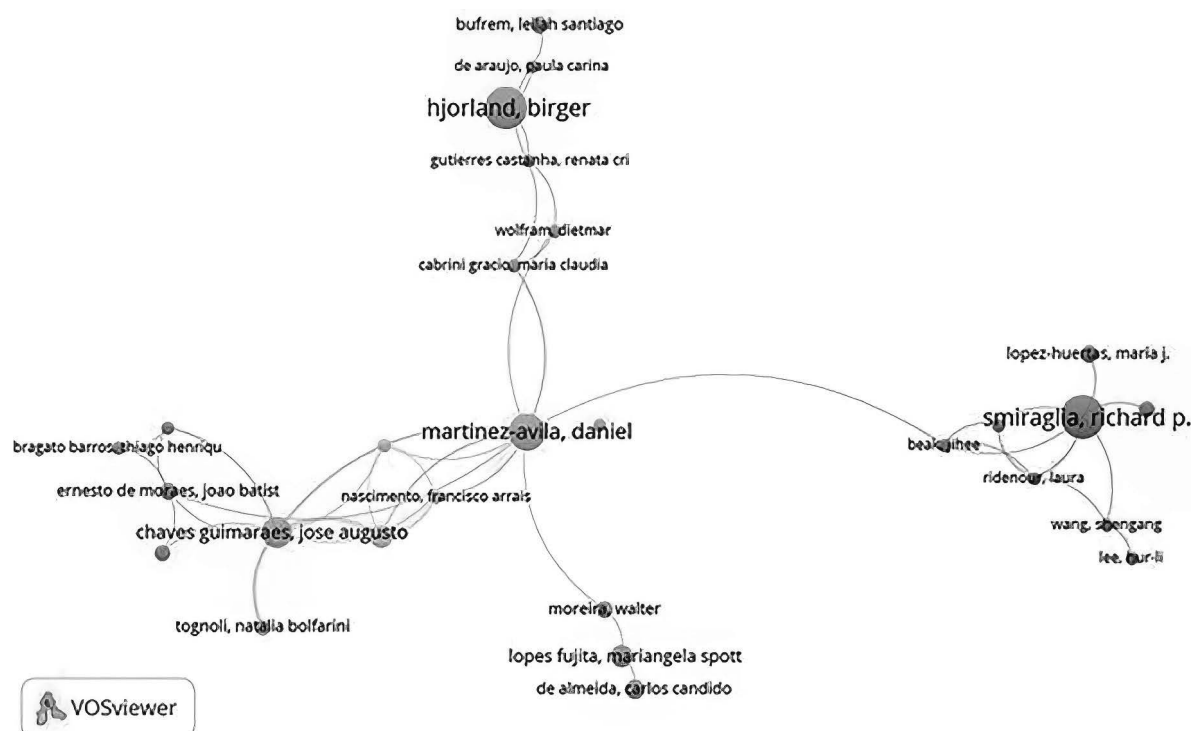


Figure 6. Co-authorship map.

thor, confirming that scientific and research papers and reports are mostly the result of the work of multiple authors (Andrés 2009), and that the academic work is increasingly the result of teamwork (Posner 2009). The network of co-authorship is shown in Figure 7.

Table 9 indicates the most prolific author is Birger Hjorland, followed by Smiraglia and Dahlberg. Also Birger Hjorland was named as the author with the most cited article, "What is Knowledge Organization (KO)?" (2008) with 93 local citations.

5.0 Conclusion

In this study, using Social Network Analysis (SNA) and scientometrics based on articles indexed in the Web of Science database on knowledge organization since the beginning, it was attempted to identify the conceptual structure of this area and to map knowledge maps, thematic trends, and authoring trends in this domain.

In order to gain a deeper understanding of the knowledge organization domain, firstly, it is necessary to identify the various research papers written in this regard and identify the underlying issues of knowledge organization by examining the subject distribution. We have identified the dominant trends in this area and analyzed research on knowledge organization from the beginning to the present.

The results of this study indicate the importance of concepts such as classification and information science that were confirmed by studying the literature on these issues.

Drawing scientific maps of knowledge organization concepts based on research conducted in this area is another result of this research. Keywords such knowledge, classification and ontology are obtained as nodes with high degrees of betweenness in the network. These nodes are very important and facilitate the formation of communication and information flow between the knowledge organization and other domains. Concepts such as knowledge, classification, ontology, library, framework, information science, epistemology and Information retrieval have the highest degree of centrality in word co-occurrence network.

Investigating the word co-occurrence network shows that the network has relatively low cohesion due to its density of 0.108, and the keywords knowledge organization, knowledge, classification as the most important co-occurring network hubs are highly connected to the other nodes, and facilitate communication with other nodes.

The findings also showed that "Nomos Verlagsgesellschaft" published the largest number of articles in this field.

Analyses such as word co-occurrence analysis are able to answer questions such as: which topics are more focused on by the scientific community? What are the different fields and subfields of science? And what evolution have they

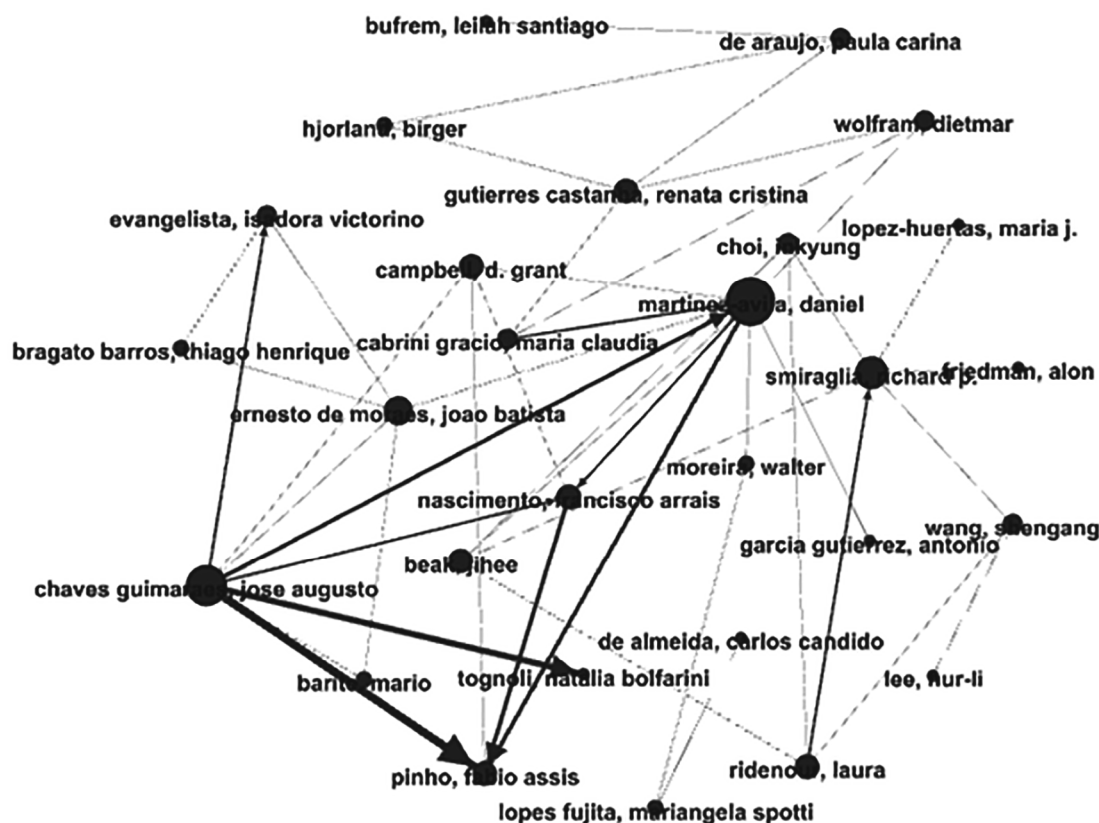


Figure 7. Map of co-authorship communications using Gephi software.

Author	Documents
Hjorland B	33
Smiraglia RP	31
Dahlberg I	28
Martinez-Avila D	21
Guimaraes Jac	17
Gnoli C	15
Tennis J	10

Table 9. Prolific authors in the field.

gone through? And what topics will probably be in the forefront of scientists' attention in the near future? The results of this study show that knowledge organization is closely related to important categories such as knowledge organization systems, information science, domain analysis, categorization, information organization, ontology and information retrieval.

Although social network analysis (SNA) research itself does not provide suggestions, it can be helpful in understanding the current situation and guiding scientific trends in knowledge organization. The present study helps re-

searchers to explain the process of studies and policy making in this area based on the identified areas of influence.

References

- Aistleitner, Matthias, Jakob Kapeller and Stefan Steinerberger. 2018. "The Power of Scientometrics and the Development of Economics." *Journal of Economic Issues* 52: 816-34.
- Andrés, A. 2009. *Measuring Academic Research: How to Undertake a Bibliometric Study*. Oxford: Chandos.

- Anghel, I. 2008. "Intellectual Capital and Intangible Assets Analysis and Valuation." *Theoretical and Applied Economics* 3: 75-84.
- Bhatia, A., and K. Aggarwal. 2018. "Impact of Investment in Intangible Assets on Corporate Performance in India." *International Journal of Law and Management* 60: 1058-73.
- Boekestein, B. 2006. "The Relation Between Intellectual Capital and Intangible Assets of Pharmaceutical Companies." *Journal of Intellectual Capital* 7: 241-53.
- Bookstein, A. 1994. "Scientometrics: New Opportunities." *Scientometrics* 30: 455-60.
- Bródka, P., K. Skibicki, P. Kazienko and K. Musiał. 2011. "A Degree Centrality In Multi-Layered Social Network." *2011 International Conference on Computational Aspects of Social Networks (CASoN)*, 237-42. doi: 10.1109/CA-SON.2011.6085951
- Broughton, Vanda, Joacim Hansson, Birger Hjørland and Maria J. Lopez-Huertas. 2005. "Knowledge Organization." Chapter 7 in *European Curriculum Reflections on Library and Information Science Education*, edited by Leif Kajberg and Leif Lorrington. Copenhagen: Royal School of Library and Information Science, 133-48.
- Buch-Hansen, Hubert. 2013. "Social Network Analysis and Critical Realism." *Journal for the Theory of Social Behaviour* 44: 306-25. doi.org/10.1111/jtsb.12044
- Castanha, R. C. G. and D. Wolfram. 2018. "The Domain of Knowledge Organization: A Bibliometric Analysis of Prolific Authors and Their Intellectual Space." *Knowledge Organization* 45: 13-22.
- Castilla-Polo, F. And C. Ruiz-Rodríguez. 2017. "Content Analysis Within Intangible Assets Disclosure: A Structured Literature Review." *Journal of Intellectual Capital* 18: 506-43.
- Chen, L. 2012. *A Mixed Methods Study Investigating Intangibles in The Banking Sector*. Ph.D. thesis University of Glasgow. <https://theses.gla.ac.uk/3267/>
- Chen, Y.-S. 2008. "The Positive Effect of Green Intellectual Capital on Competitive Advantages of Firms." *Journal of Business Ethics* 77: 271-86.
- Cheong, F. and B. J. Corbitt. 2009. "A Social Network Analysis of the Co-Authorship Network of the Pacific Asia Conference on Information Systems from 1993 to 2008." In *Proceedings. Pacific Asia Conference on Information Systems, PACIS 2009, Hyderabad, India, July 10-12, 2009*. <https://aisel.aisnet.org/pacis2009/23/>
- Cohen, M. B. 2009. *Hedonistic Pricing Models and The Valuation of Intangible Assets*. PhD thesis, Southern Cross University, Lismore.
- Crucitti, P., V. Latora and S. Porta. 2006. "Centrality in Networks of Urban Streets." *Chaos: An Interdisciplinary Journal of Nonlinear Science* 16(1): 015113. doi: 10.1063/1.2150162
- Dahlberg, I. 1993. "Knowledge Organization: Its Scope and Possibilities." *Knowledge Organization* 20: 211-22.
- Degenne, A. and M. Forsé. 1999. *Introducing Social Networks*. London: Sage.
- Edelstein, Herbert A. 1999. *Introduction to Data Mining and Knowledge Discovery*. 3rd ed. Two Crows Corporation. twocrows.com/intro-dm.pdf
- Estrada, E. and J. A. Rodriguez-Velazquez. 2005. "Subgraph Centrality in Complex Networks." *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics* 71(5 Pt 2): 056103. <https://doi.org/10.1103/PhysRevE.71.056103>
- Faust, K. 2006. "Comparing Social Networks: Size, Density, and Local Structure." *Metodoloski Zvezki* 3: 185.
- Frank, O. 2002. "Using Centrality Modeling in Network Surveys." *Social Networks* 24:385-94.
- Franks, D. W., J. Noble, P. Kaufmann and S. Stagl. 2008. "Extremism Propagation in Social Networks with Hubs." *Adaptive Behavior* 16: 264-74.
- Giju, G. C., L. Badea, V. R. L. Ruiz and D. N. Peña. 2010. "Knowledge Management - the Key Resource in the Knowledge Economy." *Theoretical & Applied Economics* 17(6): 27-36.
- Gouvêa Meireles, M. R., B. Valadares Cendón and P. E. M. De Almeida. 2014. "Bibliometric Knowledge Organization: A Domain Analytic Method Using Artificial Neural Networks." *Knowledge Organization* 41: 145-59.
- Guimarães, J. A. C. 2014. "Análise de Domínio Como Perspectiva Metodológica em Organização da Informação." *Ciência da Informação* 43(1). <https://doi.org/10.18225/ci.inf.v43i1.1415>
- Guimarães, J. A. C., J. C. Fernández-Molina, F. A. Pinho and S. O. Milani. 2008. "Ethics in the Knowledge Organization Environment: An Overview of Values and Problems in the LIS Literature." In *Culture and Identity in Knowledge Organization: Proceedings of the 10th International ISKO Conference, Montreal, Canada*, edited by Clement Arsenault and Joseph Tennis. *Advances in Knowledge Organization* 11: 361-6.
- Gumelar, M. and A. Herwany. 2018. "Benchmarking Intangible Assets in the Water Sector: an Evidence from Indonesia." *Montenegrin Journal of Economics* 14: 155-61.
- Guns, R., Y.X. Liu and D. Mahbuba. 2011. "Q-Measures and Betweenness Centrality in a Collaboration Network: a Case Study of the Field of Informetrics." *Scientometrics* 87: 133-47.
- Hall, A. D. 1969. "3-Dimensional Morphology of Systems Engineering." *IEEE Transactions on Systems Science and Cybernetics* 5: 156-60. doi:10.1109/TSSC.1969.300208.
- Hamidi, A., A. Asnafi and F. Asareh. 2008. "Analytical Survey and Mapping Structure of Scientific Publications in The Bibliometrics, Scientometrics, Infometrics and Webometrics Fields in Web of Science Database During

- 1990-2005." *Library and Information Science* 11: 161-82.
- Harris, R., R. McAdam, I. McCausland and R. Reid. 2013. "Knowledge Management as a Source of Innovation and Competitive Advantage for SMEs in Peripheral Regions." *International Journal of Entrepreneurship and Innovation* 14: 49-61.
- Hjorland, B. 2013. "Citation Analysis: A Social and Dynamic Approach to Knowledge Organization." *Information Processing & Management* 49: 1313-25.
- Hjorland, Birger. 2016. "Knowledge Organization." *Knowledge Organization* 43: 475-84. Also available in *ISKO Encyclopedia of Knowledge Organization*, edited by Birger Hjorland and Claudio Gnoli. <http://www.isko.org/cyclo/knowledgeorganization>
- Hodge, G. M. 2000. *Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files*. Digital Library Federation. <https://www.clir.org/pubs/reports/pub91/>
- Hu, C.-P., J.-M. Hu, S.-L. Deng and Y. Liu. 2013. "A Co-Word Analysis of Library and Information Science in China." *Scientometrics* 97: 369-82.
- Huang, H.-C., M.-C. Lai and T.-H. Lin. 2011. "Aligning Intangible Assets to Innovation in Biopharmaceutical Industry." *Expert Systems with Applications* 38: 3827-34.
- Jain, A. K., M. N. Murty and P. J. Flynn. 1999. "Data Clustering: A Review." *ACM Computing Surveys (CSUR)* 31: 264-323.
- Keong Choong, K. 2008. "Intellectual Capital: Definitions, Categorization and Reporting Models." *Journal of Intellectual Capital* 9: 609-38.
- Kohler, H.-P., J. R. Behrman and S. C. Watkins. 2001. "The Density of Social Networks and Fertility Decisions: Evidence from South Nyanza District, Kenya." *Demography* 38: 43-58.
- Kramer, J.-P., E. Marinelli, S. Iammarino and J. R. Diez. 2011. "Intangible Assets as Drivers of Innovation: Empirical Evidence on Multinational Enterprises in German and UK Regional Systems of Innovation." *Technovation* 31: 447-58.
- Lane, J. 2010. "Let's Make Science Metrics More Scientific." *Nature* 464: 488-9.
- Lev, B., B. Sarath and T. Sougiannis. 2005. "R&D Reporting Biases and Their Consequences." *Contemporary Accounting Research* 22: 977-1026.
- Leydesdorff, Loet and Milojevic Stasa. 2015. "Scientometrics." *International Encyclopedia of Social and Behavioral Sciences*. 2nd ed., edited by James D. Wright. Elsevier 322-7. doi.org/10.1016/B978-0-08-097086-8.85030-8
- Liu, G.-Y., J.-M. Hu and H.-L. Wang. 2012. "A Co-Word Analysis of Digital Library Field in China." *Scientometrics* 91: 203-17.
- Macías-Chapula, César A. 2004. "Bibliometric and Webometric Analysis of Health System Reforms in Latin America and the Caribbean." *Scientometrics* 53: 407-27.
- Martínez-Ávila, D. 2015. "Knowledge Organization in the Intersection with Information Technologies." *Knowledge Organization* 42: 486-98.
- Martínez-Ávila, D., J. A. C. Guimarães, F. A. Pinho and M. J. Fox. 2015. "The Representation of Ethics and Knowledge Organization in the WoS And LISTA Databases." *Knowledge Organization* 42: 269-75.
- Massa, S., and S. Testa. 2009. "A Knowledge Management Approach to Organizational Competitive Advantage: Evidence from the Food Sector." *European Management Journal* 27: 129-41.
- Miguel, S., Z. Chinchilla-Rodríguez, C. M. González and F. D. Moya Anegón. 2012. "Analysis and Visualization of the Dynamics of Research Groups in Terms of Projects and Co-Authored Publications: A Case Study of Library and Information Science in Argentina." *Information Research* 17: 524. <http://hdl.handle.net/10261/72480>
- Miyagawa, T. and S. Hisa. 2013. "Measurement of Intangible Investment by Industry and Economic Growth in Japan." *Public Policy Review* 9: 405-32.
- Mooghali A, R. Alijani, N. Karami and A. Khasseh. 2011. "Scientometric Analysis of the Scientometric Literature." *International Journal of Information Science and Manage.* 9: 19-31.
- Mrvar, A. and V. Batagelj. 2016. "Analysis and Visualization of Large Networks with Program Package Pajek." *Complex Adaptive Systems Modeling* 4: 6. doi.org/10.1186/s40294-016-0017-8
- Nakamura, L. I. 2010. "Intangible Assets and National Income Accounting." *Review of Income and Wealth* 56: S135-55. doi.org/10.1111/j.1475-4991.2010.00390
- Newman, M. E. 2005. "A Measure of Betweenness Centrality Based on Random Walks." *Social Networks* 27: 39-54.
- Okamoto, K., W. Chen and X.-Y. Li. 2008. "Ranking of Closeness Centrality for Large-Scale Social Networks." In *International Workshop on Frontiers in Algorithmics*, edited by F. P. Preparata, X. Wu and J. Yin. Lecture Notes in Computer Science 5059. Berlin; Heidelberg: Springer, 186-95. doi.org/10.1007/978-3-540-69311-6_21
- Omran, M. G., A. P. Engelbrecht and A. Salman. 2007. "An Overview of Clustering Methods." *Intelligent Data Analysis* 11: 583-605.
- Opsahl, T., F. Agneessens and J. Skvoretz. 2010. "Node Centrality in Weighted Networks: Generalizing Degree and Shortest Paths." *Social Networks* 32: 245-51.
- Posner, R. A. 2009. *Public Intellectuals: A Study of Decline*. Cambridge, MA: Harvard University Press.
- Roland, Gary and Glen Kernick. 2014. "Cost of Capital for Fair Value Reporting of Intangible Assets." In *Cost of Capital*, edited by Shannon P Pratt and Roger J Grabowski.

- Wiley, 757-78. <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118846780.ch30>
- Scott, John and Peter. J. Carrington (Eds.). 2011. *The SAGE Handbook of Social Network Analysis*. London: SAGE.
- Smiraglia, R. P. 2002. "The Progress of Theory in Knowledge Organization." *Library Trends* 50: 330-49. <http://hdl.handle.net/2142/8414>
- Steenkamp, N. and V. Kashyap. 2010. "Importance and Contribution of Intangible Assets: SME Managers' Perceptions." *Journal of Intellectual Capital* 11: 368-90.
- Sentinel Visualizer. 2009. *Social Network Analysis (SNA)*. <https://sentinelvisualizer.com/SocialNetworkAnalysis/>
- Van Raan, A. 1997. "Scientometrics: State-of-the-Art." *Scientometrics* 38: 205-18.
- Volkov, D. and T. Garanina. 2007. "Intangible Assets: Importance in the Knowledge-Based Economy and the Role in Value Creation of a Company." *Electronic Journal of Knowledge Management* 5: 539-50.
- Wang, S. 2019. "The Intellectual Landscape of the Domain of Culture and Ethics in Knowledge Organization: An Analysis of Influential Authors and Works." *Cataloging & Classification Quarterly* 57: 227-43.
- Zhao, D. and A. Strotmann. 2008. "Evolution of Research Activities and Intellectual Influences in Information Science 1996-2005: Introducing Author Bibliographic-Coupling Analysis." *Journal of the American Society for Information Science and Technology* 59: 2070-86.