Scientometrics (2019) 119:1455–1486
https://doi.org/10.1007/s11192-019-03103-1

A hierarchical approach to analyzing knowledge integration between two fields—a case study on medical informatics and computer science

Zhichao Ba[1] · Yujie Cao[1] · Jin Mao[2] · Gang Li[2]

Received: 28 July 2018 / Published online: 10 May 2019

Abstract
As a driver in modern science, interdisciplinary research has attracted a lot of attention. Major foci are laid on exploring the relations of multiple involved disciplines as well as the knowledge structure in interdisciplinary field. However, there is still a lack of decomposing the knowledge structure of interdisciplinary field to investigate how knowledge from relevant disciplines is integrated in the field. This study proposes an approach to investigating knowledge integration relationships between two research fields from a perspective of hierarchy. Medical Informatics (MI) and its most relevant field of Computer Science (CS) are chosen in the case study. This study decomposed each keyword network of the two fields into four layers by using the K-core method, then quantified the knowledge integration relationships between different layers of the two fields together. The results present that the MI basic layer shows the strongest knowledge integration with CS, followed by the middle layer, with the detail layer the weakest. And all MI layers have the greatest breadth and strength of knowledge integration with the CS middle layer, followed by the CS marginal layer and detail layer, but with the CS basic layer the weakest. A time series analysis shows that the integration of new CS knowledge into MI is a gradual process without explosive growth and the path of knowledge integration between the two fields were identified. The proposed approach could be applied to deeply understanding the integration of one discipline knowledge by an interdisciplinary field.

Keywords Interdisciplinary research · Knowledge network · Hierarchical structure · K-core · Knowledge layers

✉ Yujie Cao
cathy0021@163.com

✉ Jin Mao
danveno@163.com

Zhichao Ba
bazhichaoty@126.com

Gang Li
ligang@whu.edu.cn

[1] School of Information Management, Wuhan University, Wuhan, China

[2] Center for Studies of Information Resources, Wuhan University, Wuhan, China

Mathematics Subject Classification 05C38 · 68T30

Introduction

Interdisciplinary research (IDR), involving knowledge input from multiple disciplines, is important for scientific research. Interdisciplinary research is not only critical for academic innovation, but also essential to facilitate the development of modern science (Valentin et al. 2016). Significant breakthroughs, new discipline growth points and emerging disciplines are typically produced from interdisciplinary fields in modern science. The survey conducted in 2017 shows that 49.07% of Nobel Prize research results were achieved based on interdisciplinary collaboration (Li et al. 2017). The inherent multidisciplinarity of research problems and common research themes across many research fields are the two major reasons of flourishing interdisciplinary research (Braun and Schubert 2003).

The essence of analyzing interdisciplinary research is to observe how interdisciplinary field interacts with other related disciplines. Porter et al. (2007) considers that one notable aspect about the interdisciplinarity of a field is the integration of knowledge from other disciplines into the field, which can be measured by examining the references of the papers in the field. This method of evaluating interdisciplinarity quantifies the knowledge spread from other disciplines to the interdisciplinary field by counting the frequency of references. Due to its easy operation, this method is well applied (Karunan et al. 2017). However, it only observes the knowledge relation between interdisciplinary field and its source disciplines in a broad view of disciplines rather than in a finer-granularity level. So are a few other interdisciplinarity indicators, such as Gini coefficient and Rao-Stirling (Rafols and Meyer 2010). Meanwhile, citations are the major metaphor to observe knowledge integration among disciplines.

A few recent studies have attempted to distill the knowledge structure of interdisciplinary field to observe how it develops (Lee and Jeong 2008; Liu et al. 2012; Piepenbrink and Nurmammadov 2015). Identifying topics of a field is a well applied approach by either co-word analysis (Besselaar and Heimeriks 2006) or topic modeling (Song and Kim 2013). Detected topics may show apparent characteristics of a specific discipline. With this idea, Nichols (2014) proposed a topic-level Stirling indicator to measure the interdisciplinarity of proposals at the National Science Foundation. Further, Xu et al. (2016) looks at the level of terms and proposed an indicator to measure the interdisciplinarity of terms. Compared with discipline-level methods and topic-level methods, this method observes the knowledge associations between interdisciplinary field and related disciplines at a much finer granularity. Specific topics, theories, and methods of related disciplines could be discovered and analyzed by examining the content of the field. In the study (Xu et al. 2016), the involved disciplines and their specific involved topics were identified and analyzed in a long-time span for the field of Information Science and Library Science. To deeply understand the development of interdisciplinary field, more similar approaches are needed to uncover detailed knowledge integration relations between an interdisciplinary field and its related disciplines.

In this paper, we propose a new approach to the measurement of the knowledge integration and its evolution among two fields from the perspective of hierarchy. Representing the specific concepts in a direct way (Wang et al. 2013), keywords of articles are considered as the most basic fundamental carrier of knowledge (Lee et al. 2010) and are well representative for articles' ideas and topics (Griffiths and Steyvers 2004). Deferring to these ideas, we also look at the knowledge system of a field through the terms. Co-word network is used to reveal the keyword correlation patterns acknowledged by the research community in the field. In our study, we recognize that the co-occurrence of the two terms in a research paper manifests the integration of the knowledge carried by the two terms in the paper. With this premise, knowledge integration between two disciplines could be analyzed by investigating on the co-word network consisting of terms from the two disciplines. In this study, we take Medical Informatics (MI) as a case of interdisciplinary discipline (Morris and McCain 1998). And, Computer Science (CS) is one of its highly related disciplines in terms of references. In addition, the K-core decomposition method has been applied to obtain hierarchical layers of the knowledge from co-word network (Xiao et al. 2016). The hierarchical layers of the two disciplines will be obtained by this method so that the knowledge integration between the layers of the two disciplines will be investigated. Specifically, this study attempts to address the following three research questions:

RQ#1 What are the hierarchical layers of Medical Informatics and Computer Science consisting of terms? Are there any shared terms between them?
RQ#2 How is the knowledge in different layers of Computer Science integrated by different knowledge layers of Medical Informatics?
RQ#3 How has the integration of knowledge from Computer Science into Medical Informatics been changing over time?

This study offers a microscopic perspective for understanding the hierarchical structure and evolutionary characteristics of interdisciplinary knowledge interactions. Our contributions are multifold. First, this study proposes a generalized approach to analyzing knowledge integration relationships between two fields through the decomposed knowledge network layers of the two fields. The number of shared terms and their connected edges in the knowledge network are used to quantify the breadth and strength of knowledge integration. The proposed approach could be applied to deeply understanding knowledge integration between two fields, especially for the case that an interdisciplinary field integrates the knowledge from another field. Second, the knowledge integration of CS knowledge by MI was examined in detail, which reveals the knowledge integration process and path of CS by MI. The findings of this study can be used to master the frontier dynamics of interdisciplinary field and guide the development direction of discipline.

Related work

Related research on knowledge relationship between interdisciplinary field and related disciplines could be found in the following two research directions, the evaluation of interdisciplinarity and the analysis of knowledge structure.

The evaluation of interdisciplinarity

One major focus of studying interdisciplinary field is to manifest its interdisciplinarity by exploring the evidence that multiple disciplines take their effects in the field. In other word, the knowledge from other disciplines is investigated in the field. To this end, determining the disciplines of knowledge items, such as journals, articles, keywords, and so on, is the critical prerequisite. Some scholarly databases such as Web of Science and Scopus assign subject categories to the indexed journals or articles, which facilitates subsequent discipline-related analysis. Currently, a large amount of scientometrics studies on interdisciplinary research have taken this advantage (Chi and Young 2013; Leydesdorff and Rafols 2011). For example, Morillo et al. (2001) employed the co-classification method to study interdisciplinarity based on the subject categories of journals, finding that multi-assigned journals don't show higher interdisciplinarity than single-assigned journals. Karunan et al. (2017) used co-citation analysis to discover interdisciplinary interactions by observing common papers and cross-disciplinary citations. The discipline information of other knowledge items could be inferred based on their associations with the articles. Xu et al. (2016) inferred the discipline information of terms according to the subject categories of articles together with term frequency and the betweenness value on co-occurrence network. The affiliations of authors provide another source to reflect the disciplines of authors, which is often applied to investigate interdisciplinary collaboration (Karlovčec and Mladenić 2015; Abramo et al. 2012; Bergmann et al. 2016). This method considers social aspect and research practice of interdisciplinarity rather than the attributes of papers (Schummer 2004).

With the discipline information of knowledge items, a multitude of bibliometric indicators have been proposed to measure the interdisciplinarity of a research field. Earlier indicators include Citation Outside Category (Porter and Chubin 1985), Salton Index (Salton and Mcgill 1988), Disciplinary Impact Factor (Hirst 1978), Journal to Field Impact Score (Leeuwen and Moed 2005), and so on. Researchers have attempted to conclude the properties of interdisciplinarity. Stirling (2007) provided a general framework for analyzing the diversity of interdisciplinarity by using the variety, balance and disparity indicators. Rafols and Meyer (2010) further captured the interdisciplinarity of bionanoscience by network coherence and diversity indicator. Afterwards, Leydesdorff and Rafols (2011) investigated betweenness centrality, Gini coefficient and Rao-Stirling measures for interdisciplinarity. They argued that Shannon entropy performs better than the Gini coefficient among the vector-based indicators, but is sensitive to size. Porter et al. (2007) derived three measures based on ISI Subject Category information for sets of research papers–Integration (I), Reach (R) and Specialization (S) to gauge the interdisciplinarity. In summary, the diversity and heterogeneity of interdisciplinarity make traditional one-dimensional indicators unsuitable for measuring the scope, pattern and dynamic mechanism of IDR. The practice proves that some indicators are subjective and hard to conduct, and lack of contrastive analysis to explain the validity of each indicator. A slice of standardized, multi-dimensional and systematic quantitative indicators remain further to be developed.

In addition, an alternative way is to apply text mining to discover interdisciplinary correlations of knowledge. Nanni et al. (2017) investigated the usefulness of different text mining for identifying interdisciplinary practices directly from the textual content, finding that word-features are consistently outperforming topic-features from LDA. Nichols (2014) provided a topic model approach to measure interdisciplinarity at the National Science Foundation (NSF) by assessing the language and content of award proposals. He pointed that co-funding or co-review relationships could provide a means of measuring interdisciplinarity in NSF portfolios as well.

The detection and analysis of knowledge structure of interdisciplinary field

The cognitive structure of interdisciplinary field is the fruit of interdisciplinary research practice, which could be used to manifest the characteristics of IDR. The internal and external scholarly interactions of a field can be examined by mapping the publications in the field (White and McCain 1997). The principal methods, which are also well used for interdisciplinary fields, are co-word analysis, bibliographic coupling, co-citation analysis, and inter-citation analysis (Liu and Wang 2005). These methods could be treated as network-based methods in that the relations of knowledge items are examined by these methods. The interaction relationships and boundary of interdisciplinary field can be identified, and even their dynamic evolutions over time can be revealed. Hu and Zhang (2017) built co-occurrence network of disciplines to analyze the structure and pattern of interdisciplinary collaborations in Big Data research. Chi and Young (2013) utilized journal citation analysis to discover the collaboration of major disciplines in a research domain, and the intercultural relations over time. Meanwhile, dynamic knowledge integration can be explained by revealing the knowledge interaction patterns through co-word analysis over time (Wang et al. 2013). However, Wang et al. (2013) only measured the integration degree among the fields, but did not reveal the detailed integration process, which is our attempt in this study.

On the other hand, different from the above research about the interdisciplinarity on the macroscopic view, the internal structural properties of the interdisciplinary field at the micro level, such as modularity, cohesion and diversity nature of internal topics or cooperative relationships, also have been extensively studied. Liu and Xia (2015) explored the cooperation structure for "evolution of cooperation" (EOC) field and found that the co-authorship network of this interdisciplinary field gradually evolved from "core-periphery" structure to modular structure. Lee et al. (2016) classified topics into the two groups of subjects and methods to discover central subjects, central methods, central subject–method pairs, and major subjects–methods groups from communication studies between 1990 and 2004. Dong et al. (2018) identified important interdisciplinary topics in Information Science and Library Science (LIS) by integrating various methods, including co-occurrence network analysis, high-TI terms analysis and burst detection method.

Hierarchy is an inherent feature of science system (Cole 1983). In the sense of semantic space, science displays with hierarchical structure, e.g., the relationship among hard science, physics and hydromechanics. Similarly, the knowledge structure of a field also demonstrates such hierarchy. Recent complex network studies have found that a variety of real networks exhibit hierarchical organization (Clauset et al. 2008). Sales-pardo et al. (2007) measured node affinity based on modularity method and proposed a box-model clustering to identify modules at hierarchical level. Lancichinetti and Fortunato (2012) found the local maxima of a fitness function to uncover the hierarchical and the overlapping community structure of complex networks. Carmi et al. (2007) decomposed the Internet topology into three components using K-shell decomposition, which provided insights into the hierarchical structure of the Internet and its functional consequences. However, the hierarchy of knowledge network has been paid little attention to, especially for interdisciplinary fields. Recently, Xiao et al. (2016) explored the topic hierarchy of the digital library field in which keyword networks were divided into four hierarchical layers.

K-core decomposition, as a network analysis tool, can highlight interesting structural properties that are not captured by degree distribution or other simple topological measures (Alvarez-Hamelin et al. 2017). This method identifies particular subsets of the network and provides a probe to explore the properties of network's regions with increasing centrality and connectedness (Zhang et al 2010). The core with larger index corresponds to the nodes having larger degree and taking more central positions in the network. K-core decomposition has been extensively used to study the hierarchical structure of large scale networks (Alvarez-Hamelin et al. 2005; Liu et al. 2015; Khaouid et al. 2015; Eidsaa and Almaas

2013). In this study, we therefore apply the K-core approach to subdivide knowledge network into hierarchical layers, and attempt to analyze knowledge integrations based on the decomposed layers.

Methodology

Data collection and preprocessing

As a representative interdisciplinary field, Medical Informatics (MI) was chosen in the case study, which cites a large amount of papers from the discipline of Computer Science (CS). In Web of Science (WoS), each journal is assigned to at least one of the subject categories. We downloaded the metadata of papers published in all the 24 journals under the category of MI in the 2016 version of WoS (Appendix A). In total, we collected 34,727 MI articles from 1990 to 2016. The references were parsed and the referenced journals were ranked according to their cited counts. We then selected 30 journals (Appendix B) to represent CS according to their category assignments.[1] A total of 69,350 CS articles in these journals from 1987 to 2016 were downloaded from WOS. The start year of CS was previous to that of MI, because we hope that the knowledge of the CS discipline could cover those CS knowledge adopted by MI.

We use both author keywords and title keywords in this study. Author keywords were parsed from the metadata of the papers. Since author keywords are often absent in a large portion of papers in a field (Mao et al. 2018), we include title keywords as the complement. Meaningful noun phrases were extracted as keywords from the titles by using part-of-speech tagging based on Stanford CoreNLP toolkit (Manning et al. 2014) and frequent pattern mining with the FP-Growth algorithm (Han et al. 2004). Author keywords and title keywords were combined and cleaned by merging acronyms and their full forms, identifying synonyms, and filtering general terms, such as "problems," "informatics," "data," and so on.

In addition, we unified some synonyms expressed differently in MI and CS. The common terms shared by the two disciplines were identified. We define these terms as "interdisciplinary terms" that reflect the common knowledge and can be seen as the bridge connecting the knowledge networks of two fields. In theory, these interdisciplinary terms belong to three groups, i.e., the knowledge originally innovated in MI, CS and other fields. However, we only focus on the integration of CS terms into MI in this study. The disciplinary affiliation of these terms were then labeled by one expert from the medical informatics field and two authors according to their own understanding, aiming at distinguishing CS terms from others.

Knowledge network construction

We constructed knowledge network of a field by using the terms as nodes and forming edges according to the term co-occurrence in the papers. The weight of an edge indicates

[1] It should be noted that the CS discipline defined by the 30 journals is a small subset of artificial intelligence field in computer science. Other fields of computer science are not covered, such as cybernetics, hardware, software engineering, etc.

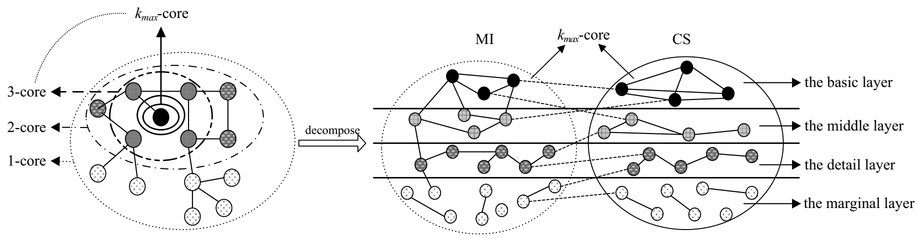


Fig. 1 Sketch of the K-core decomposition for the knowledge networks of MI and CS, whose highest core indexes ($k_{\max}$-core) could be different. Each closed ellipse contains a set of terms belonging to a given *k*-core. Colors on the nodes distinguish different *k*-shells. Dashed lines of the right chart connect the same terms in the two fields

the number of papers that the two terms connected by the edge co-occur in. The edges with weights lower than 2 were trimmed. Two knowledge networks were built for MI and CS respectively.

Knowledge network decomposition with K-core

We split the knowledge networks of CS and MI into hierarchical layers with the K-core decomposition method. In a network, the degree of a node indicates the number of connections incident to the node. A K-core of a network is a maximal connected sub-network in which the degrees of the nodes are at least *k*. The K-core (also called the shell index) of a node is defined as *k* if it belongs to the K-core but is removed in the (k+1)-core (Dorogovtsev et al. 2006). Thus, the K-core of a network can be obtained by removing recursively all the nodes with the degree less than *k*, until the degrees of all nodes in the remaining graph are at least *k* (Alvarez-Hamelin et al. 2005). Based on the procedure, the network can be decomposed into *k*-cores. We start by removing all nodes with one connection only and assign them to the 1-core. In the same manner, we recursively remove all nodes with the degree of 2 (or less), creating the 2-core. We continue by increasing *k* until all nodes in the network have been assigned to one of the cores. The highest core index is named as $k_{\max}$. Nodes in higher cores are considered to be more central to the network (Carmi et al. 2007).

It is worth to note that the K-core decomposition may produce a few similar shells which are adjacent in the hierarchical structure and could be further merged (Xiao et al. 2016; Carmi et al. 2007). By applying the idea of "first decomposing and then merging", we partitioned a knowledge network into four layers by combining similar shells into the same layer. Clustering coefficient (Collins and Chow 1998) was used to measure the similarity between adjacent shells. We obtained the cluster coefficient of a shell by averaging the clustering coefficients of all nodes in the shell. By using this method, the terms in both MI and CS knowledge networks were divided into the four kinds of layers respectively, as illustrated in Fig. 1. It should be mentioned that the shared terms by the two disciplines could belong to different kinds of layers in the two disciplines.

Quantifying the knowledge integration of CS into MI

The integration of CS knowledge by the MI field could result in certain structural characteristics of the MI knowledge network. The co-occurrence of two terms is assumed as the metaphor for the integration of the knowledge conveyed by the two terms. Thus, how CS

Table 1 Basic structure properties for the knowledge networks of MI and CS

Knowledge network	Size	Edge	Density	Clustering coefficient
MI	1043	2662	0.013	0.389
CS	1252	3033	0.014	0.957

knowledge is integrated into MI could be analyzed by examining the nodes and edges of the MI knowledge network that connect with the CS terms.

We quantified the knowledge integration from two aspects. One is the CS terms connected by MI, which reflects the extent to which CS knowledge is integrated by MI. The other is the nodes and edges linked with the CS terms, which indicate the breadth and strength of knowledge integration for these terms. We further measured the breadth of knowledge integration by the number of the MI terms that the CS terms connect with and the strength of knowledge integration by the weight of the edges connecting the CS terms. For both metrics, the intra-connections among the CS terms in the MI network are not taken into consideration. With the layers discovered for CS and MI, we analyzed the knowledge integrations between CS and MI layers. These metrics reflect the closeness of the knowledge relationship between the two fields. What's more, the evolution of their knowledge integrations over time was then analyzed.

Results

The knowledge networks of MI and CS

For the final knowledge networks, 1043 keywords of MI and 1252 keywords of CS were retained. Table 1 shows the basic structure properties of the two knowledge networks. The two fields share 277 common keywords. The Kappa coefficient (Cohen 1960) for testing the discipline categorization of keywords achieved an agreement of 0.89, showing a high consistency among the three annotators. After reaching agreement through discussion among the authors, 105 CS terms were labeled from the shared keywords (Appendix C).

The hierarchical layers in the knowledge networks of MI and CS

By the K-core decomposition, the knowledge network of MI was divided into 8 cores as shown in Fig. 2, and the knowledge network of CS was divided into 9 cores as shown in Fig. 3.

The structural characteristics of each core in Table 2 show that as the core index K_c increases, the average node degree and density of the core gradually increase, but the number of nodes in the core decreases. The greater the average degree of nodes in the core is, the higher the core index is. Along with the increase of K_c, both p_1 and p_2 achieve the maximum in $K_{\max}$-core (8th and 9th core for MI and CS respectively). The connections with $K_{\max}$-core to each core generally account for a large proportion of all edges connected by the core ($\bar{p}_3 = 36.68\%$ for MI and 46.68% for CS), whereas the proportion of internal connections in the core are small ($\bar{p}_1 = 2.47\%$ for MI and 1.93% for CS). It is clear that the highest core has a great impact on other cores.

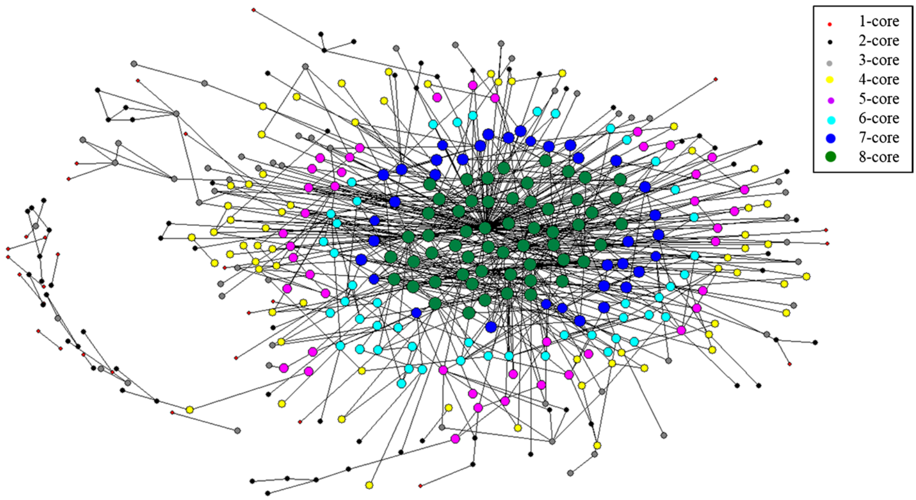


Fig. 2 The 8 cores of the MI knowledge network. The network is simplified for visualization by selecting the edges with weights greater than 2

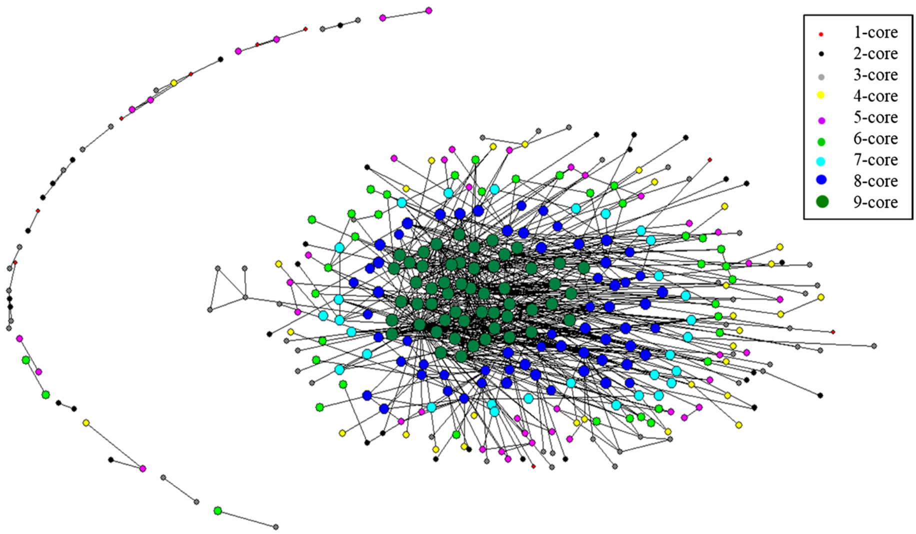


Fig. 3 The 9 cores of the CS knowledge network. The network is simplified for visualization by selecting the edges with weights greater than 5

The fluctuation of clustering coefficient in each core is significant (Table 2), which means that clustering coefficient is more appropriate to depict the hierarchical structure of the network than other metrics like average node degree and network density (Xiao et al. 2016). In MI, the clustering coefficients of 3-core and 8-core are exceedingly large, while the clustering coefficients of other cores are much smaller. Using similar methods by Xiao et al. (2016) and Carmi et al. (2007), we merged these cores into four layers by taking

Table 2 The structure characteristics of each core in MI and CS

K_c-core	MI								CS								
	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	9
$N(K_c)$	**372**	221	147	88	58	58	41	58	**522**	285	154	84	40	40	51	41	35
N_p/%	**35.67**	21.19	14.99	8.44	5.56	5.56	3.93	55.61	**41.69**	22.76	12.30	6.71	3.19	3.19	4.07	3.27	2.79
Edge	41	100	104	50	44	63	58	**402**	122	101	68	40	22	36	71	70	**283**
$<C>$	0	0.19	**0.40**	0.18	0.17	0.17	0.16	0.36	0	0.47	0.54	0.21	0.25	0.29	0.32	0.29	**0.60**
$d(K_c)$	0	0	0.01	0.01	0.03	0.04	0.07	**0.24**	0	0	0	0.01	0.03	0.05	0.06	0.08	**0.48**
$<k>$	0.16	0.41	0.97	1.31	2.66	3.81	7.07	**24.32**	0.08	0.25	0.58	1.18	2.82	4.61	5.57	8.13	**47.56**
p_1/%	1.54	3.76	3.91	1.88	1.65	2.37	2.18	**15.10**	2.01	3.33	2.24	1.32	0.73	1.19	2.34	2.31	**9.33**
p_2/%	19.62	23.04	19.81	11.71	10.73	12.28	10.57	**27.92**	15.87	17.67	13.36	9.37	8.10	10.39	11.384	11.14	**17.92**
p_3/%	25.84	25.58	25.14	35.13	35.61	40.64	**42.99**	–	44.08	49.34	45.19	48.01	43.46	45.02	45.37	**52.98**	–

K_c = index of the core, $N(K_c)$ = number of nodes in the K_c-core, Np = proportion of nodes in the core against all nodes, E = number of edges in the core, $<C>$ = clustering coefficient in the core, $<k>$ = average degree of nodes in the core, $d(K_c)$ = density of the core, p_1 = proportion of edges in the core against all edges in the network, p_2 = proportion of internal edges in the core against all edges connecting with the nodes in the core, p_3 = proportion of edges connecting with the highest core against all edges connecting with the nodes in the core. Bold values indicate the highest for layers of MI or CS

3-core and 8-core as the demarcations. Likewise, 3-core and 9-core are the two demarcations dividing the CS knowledge network into four layers.

Table 3 provides descriptive statistics information about the four layers for MI and CS. The basic layer contains the K_{max}-cores of the knowledge networks, which can be viewed as the main backbone of the field since it connects with a large portion of nodes in other cores (p_3 in Table 2). Although it includes the least terms, the number of inner edges and all connected edges to this layer are more than the detail layer and the marginal layer. The middle layer contains more cores than other layers, thus it has the most edges of both internal and external. The layer is an intermediate transition region, in which the terms can be used to recognize how the general concepts in the basic layer evolve into concepts in the detail layer (Xiao et al. 2016). The detail layer connects directly with the above two layers. It reveals the microstructure of the knowledge network and can be used to discover specific emerging topics. The marginal layer contains the most nodes but is the sparsest. The number of connected edges to this layer is more than the number of inner edges in the layer, indicating the marginal layer is highly attached to the other layers. In fact, terms in this layer are either isolated pairs or peripheral terms as shown in Figs. 2 and 3.

In addition, the MI marginal layer contains the most interdisciplinary terms and CS terms, followed by the MI middle layer. The hierarchical distribution of 105 CS terms in CS knowledge network and their connections with the layers of MI are listed in Appendix C. Although only a total of 5 CS terms in the MI layer, we can confirm that MI is a fairly typical interdisciplinary field in that the MI basic layer includes 36 interdisciplinary terms accounting for 62.10% of the layer.

Interdisciplinary interactions between MI and CS

To reveal the interdisciplinary interactions between different layers of MI and CS, we mainly focus on how the CS terms from different layers have been integrated into each layer of MI. The connections between the terms in the MI layers and CS terms are visualized in the demo web page[2] for reference.

The MI basic layer

Table 4 shows the distribution of the terms and the edges from different CS layers associated with the MI basic layer. A total of 58 CS terms were connected by the MI basic layer, which occupies 55.23% of the 105 integrated CS terms. They connect with other 44 terms (75.86% of the layer) besides the 5 CS terms in this layer and only disconnected with 9 terms, reflecting a prominent breadth of knowledge integration between them. The weight of the edges connecting CS terms in this layer is great with an average of 3.07, which is much higher than the average weight of all edges in this layer (1.75). This implies that the MI basic layer has a strong strength of knowledge integration with the CS terms.

The MI basic layer integrates the least terms from the CS basic layer (Table 4). The number of nodes and the cumulative weight of the edges connected by the CS basic layer are both low, while, the average weight of edges is the largest, which means that their associations are strong. The MI basic layer integrates 20 terms from the CS middle layer, showing that the breadth and strength of knowledge integration between them are the most

[2] https://shiyun-w.github.io/MI-layer/.

Table 3 Descriptive statistics of different layers for MI and CS

	MI						CS			
	Cores	Terms	Interdisciplinary terms	CS terms	Inner edges	All connected edges	Cores	Terms	Inner edges	All connected edges
The basic layer	8	58	36	5	402	1842	9	35	283	1580
The middle layer	4, 5, 6, 7	245	82	35	498	2114	4, 5, 6, 7, 8	256	559	2269
The detail layer	3	147	34	12	104	581	2, 3	439	213	1084
The marginal layer	1,2	593	125	53	172	780	1	522	122	769

Table 4 The distribution of the terms and the edges from different CS layers associated with the MI basic layer

Layers of CS	The basic layer of MI			
	Number of CS terms	Nodes connecting CS terms	Weight of edges connecting CS terms	Top 5 CS terms with high breadth of integration
The basic layer	4	7 (12.07%)	32 (3.56)	Model selection; Machine learning; Classification; Mixture model
The middle layer	20	39 (67.24%)	286 (3.14)	Decision support system; Decision making; Social networking sites; Information system; Natural language processing
The detail layer	8	14 (24.14%)	53 (2.79)	Online system; Web site; Human computer interaction; Multi-media systems; Knowledge acquisition
The marginal layer	26	26 (44.83%)	160 (2.96)	DEA analysis; User computer interface; Document analysis; World Wide Web; FIR filter
Total	58	44 (75.86%)	531 (3.07)	

prominent. It reveals that the MI basic layer has the closest interdisciplinary interaction with the CS middle layer. Eight terms from the CS detail layer have been connected by the MI basic layer. The interaction between them is broader and stronger than the interaction between the MI basic layer and CS basic layer. Although the MI basic layer integrates the most terms from the CS marginal layer, the breadth and average strength of the integration are less than with the CS middle layer. In sum, the MI basic layer has the broadest and strongest integration relationship with the CS middle layer, followed by the CS marginal layer.

The MI middle layer

The MI middle layer integrates 60 terms from the four layers of CS, which connect with other 107 MI terms in this layer as shown in Table 5. Compared with the MI basic layer, the breadth of knowledge integration for CS terms in the MI middle layer is relatively high in terms of the number of connected nodes. However, the weight of all edges connecting CS terms in this layer is lower than in the MI basic layer, indicating that the MI middle layer has a weaker strength of knowledge integration for CS terms than the MI basic layer. In addition, the average weight of edges connecting CS terms in this layer is lower than the MI basic layer as well.

The MI middle layer only combines two general concepts, “model selection” and “classification”, from the CS basic layer. These terms just connect with 4 MI terms despite their strong edges (avg. 3.50). It means that most of terms in the CS basic layer has not yet been absorbed or applied extensively to the MI middle layer. The MI middle layer connects with 19 CS middle layer terms with the highest breadth and strength of knowledge integration among all CS layers. A total of 8 terms from the CS detail layer have been connected by the MI middle layer. However, the number of nodes and the weight of the edges connected by this layer are both low. It is noticeable that the MI middle layer connects with the most terms form the CS marginal layer where terms are less influential in CS. The breadth and strength of knowledge integration with this layer are higher than with the CS basic layer and CS detail layer. On the whole, the MI middle layer has the broadest and strongest interaction with the CS middle layer, followed by the CS marginal layer.

The MI detail layer

There are in total 28 CS terms integrated by the MI detail layer as shown in Table 6. They only connect with other 39 terms with low weight of edges. It is clear that the breadth and strength of knowledge integration for CS terms in the MI detail layer are smaller than in the MI basic layer and the MI middle layer. It’s partially due to the small scale of this layer (147 terms, Table 3). In addition, the lower average weight of the edges connecting CS terms also evidences that the strength of knowledge integration for CS terms in this layer is weak.

What is also apparent from the Table 6 is that, for the MI detail layer, the breadth and strength of knowledge integration with each layer of CS are all small. The CS basic layer has two terms, “classification” and “support vector machine,” integrated by the MI detail layer. However, they just co-occur with one MI term but connect with 4 other CS terms in this layer. It is shown that the interactions between them are the weakest among with other CS layers. Although the MI detail layer associates with the same number of terms from the CS middle layer and the CS detail layer, it has greater breadth and strength of knowledge

Table 5 The distribution of the terms from different CS layers of CS and their associations with the MI middle layer

Layers of CS	The middle layer of MI			
	Number of CS terms	Nodes connecting CS terms	Weight of edges connecting CS terms	Top 5 CS terms with high degree
The basic layer	2	4 (1.63%)	14 (3.50)	Model selection; Classification
The middle layer	19	71 (28.98%)	242 (2.75)	Decision making; Decision support system; Agility; Information system; Natural language processing
The detail layer	8	17 (6.94%)	36 (2.12)	Web site; Human computer interaction; Non linear dynamics; Multi-media systems; Online system
The marginal layer	31	56 (22.86%)	166 (2.52)	Data quality; Document analysis; DEA analysis; User computer interface; World Wide Web
Total	60	107 (43.67%)	458 (2.62)	

Table 6 The distribution of the terms from different CS layers and their associations with the MI detail layer

Layers of CS	The detail layer of MI			
	Number of CS terms	Nodes connecting CS terms	Weight of edges connecting CS terms	Top 5 CS terms with high degree
The basic layer	2	1 (0.07%)	4 (2.00)	Classification; Support vector machine
The middle layer	8	11 (7.48%)	24 (2.02)	Decision making; Decision support system; Natural language processing; Social networking sites; Agility
The detail layer	8	15 (10.20%)	40 (2.67)	Spiking neuron; Complex network; Health information technology; Distance education; Online system
The marginal layer	10	18 (12.24%)	42 (2.21)	World Wide Web; System development; DEA analysis; User computer interface; Decision analyses
Total	28	39 (26.53%)	110 (2.29)	

integration with the CS detail layer than the CS middle layer. However, the MI detail layer shows the closest interaction with the CS marginal layer due to the broadest and strongest integration relationship between them.

The MI marginal layer

The MI marginal layer integrates 45 terms from CS connecting with 72 other terms besides CS terms in this layer (Table 7). The breadth and strength of knowledge integration for CS terms in the MI marginal layer are larger than in the MI detail layer, but smaller than in the MI basic layer and the MI middle layer. Thus, combined with the above analysis, we can conclude that the MI middle layer has the closest interaction relationship for CS in terms of the breadth and strength of knowledge integration, followed by the MI basic layer and MI marginal layer, with the MI detail layer furthest.

The MI marginal layer integrates the two core terms, "machine learning" and "neural network" from the CS basic layer. In spite of the high weight of the edges, only two terms besides CS terms are linked with them in the MI marginal layer. It is evident that the breadth and strength of knowledge integration between the MI marginal layer and the CS basic layer are the smallest. Although the MI marginal layer connects with fewer terms from the CS middle layer than the CS marginal layer, the number of nodes and the weight of edges connected by the CS middle layer are bother larger, indicating that their knowledge integration are the broadest and strongest. In addition, the MI marginal layer has a stronger integration with the CS marginal layer than with the CS detail layer and the CS basic layer. Overall, the MI marginal layer has the closest interaction with the CS middle layer, followed by the CS marginal layer.

The evolution of knowledge integration between MI and CS

Overall trend

We analyzed the trend of integrating CS into MI retrospectively. Figure 4a shows the dynamic trends of the number of CS terms yearly integrated by the MI layers. It is observed that each MI layer has different growth rate of integrating CS terms over time. The growth is remarkable for the MI basic layer, the MI middle layer and MI marginal layer, but remains relatively stable for the MI detail layer. From a dynamic perspective, the MI middle layer and the MI basic layer integrated the most CS terms every year, followed by the MI marginal layer, and the least is the MI detail layer. The results are in accordance with the overall integration of CS terms by the MI layers analyzed in the section of *Interdisciplinary interactions between MI and CS*. In addition, the trends of the number of terms from different CS layers were integrated by MI are also different when MI is considered as a whole, as shown in Fig. 4b. Less CS basic layer and CS detail layer terms were integrated than the CS middle layer and CS marginal layer. In sum, the number of the CS terms integrated by MI is gradually increased over time.

Knowledge integration between layers

We further investigate the knowledge integration layer by layer to identify the integration process and path between the two fields. Figure 5 shows the number of the terms from each CS layer integrated into the MI layers at each of the selected years.

Table 7 The distribution of the terms from different CS layers and their associations with the MI marginal layer

Layers of CS	The marginal layer of MI			
	Number of CS terms	Nodes connecting CS terms	Weight of edges connecting CS terms	Top 5 CS terms with high degree
The basic layer	2	2 (0.03%)	6 (3.00)	Machine learning; Neural network
The middle layer	14	28 (4.72%)	71 (2.45)	Semantic network; Decision making; Natural language processing; Social networking sites; Text mining
The detail layer	10	19 (3.20%)	53 (2.21)	Sensitivity analyses; Non linear dynamics; System Modeling; Spiking neuron; Online system
The marginal layer	19	24 (4.05%)	62 (2.83)	DEA analysis; Data analysis; World Wide Web; Data quality; Computer simulation
Total	45	72 (12.14%)	192 (2.49)	

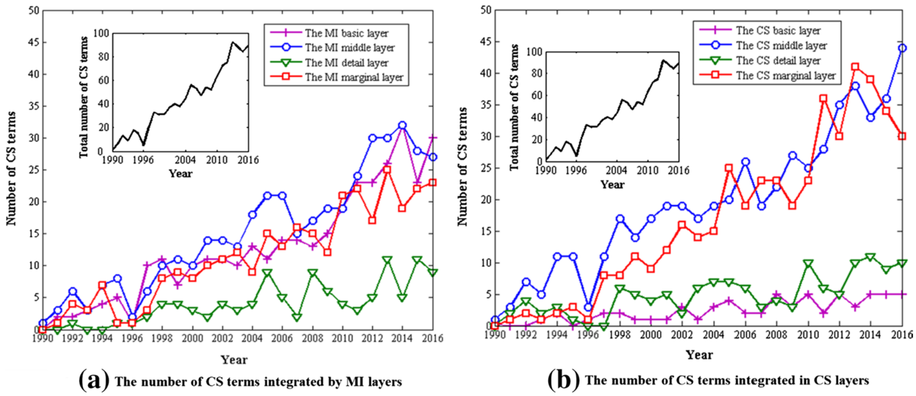


Fig. 4 The trends of the number of CS terms integrated by MI from 1987 to 2016

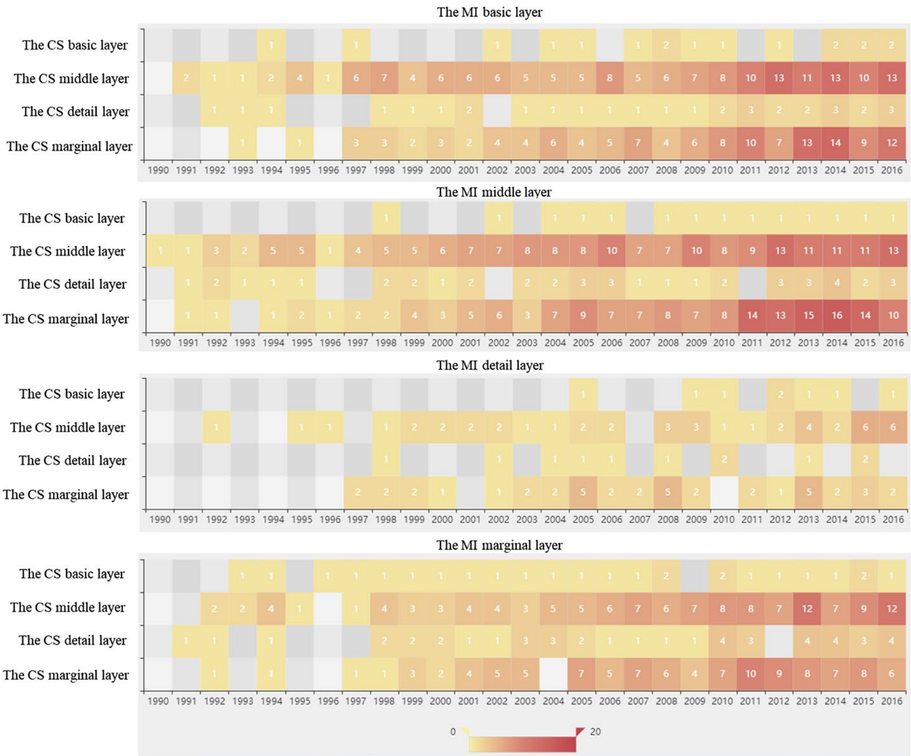


Fig. 5 The frequency distribution of the CS terms integrated by MI layers at each year from 1987 to 2016

According to Fig. 5, a growth tendency of terms in each CS layer integrated into MI over time can be observed. Each MI layer integrated only a few CS terms in the early period, and integrated more terms in the later period. This indicates that the interdisciplinary interaction between MI and CS became increasingly closer over time. As the development of MI, more information technologies from CS were applied, such as "natural language processing", "text mining", "knowledge discovery", and so on. In addition, the increasing number of MI papers published every year might result in the increasing number of integrated CS terms as well. The years that different MI layers firstly integrated CS terms are different. In 1990, only one term, "decision support system" from the CS middle layer was integrated into the MI middle layer. The MI basic layer and the MI marginal layer firstly integrated CS terms in 1991, whereas the MI detail layer in 1992.

Terms from the CS middle layer and the CS marginal layer were integrated more frequently than terms from the CS basic and detail layers. As shown in Appendix C, the terms with largest integration breadth or strength are from the CS middle layer. Those middle layer terms are either problem-oriented like "decision making", or representing a group of techniques, such as "decision support system" and "natural language processing". Many marginal terms are specific techniques that are well applied in medical informatics, e.g., DEA analysis, or CS related problems that MI shows interested in, e.g., user computer interface and data quality. It's interesting that basic terms from CS, such as "machine learning" and "neural network", are not the most frequently integrated terms. The probable reason could due to that researchers typically prefer to state semantically narrow terms of specific techniques and problems in the papers.

A significant ascending trend can be observed in terms of the number of integrated CS terms every year. From the latter two layers, only 3 to 5 terms were combined into each MI layer at most years. In particular, terms integrated before 1998 were mainly from the CS middle layer, and they were mostly tied with the MI basic and middle layers. From 1998 on, a larger number of terms from the CS marginal layer were integrated by MI layers, except for the MI detail layer.

Yearly applied new CS terms by MI

To further analyze critical time, the years when 105 CS terms were firstly integrated are examined in Fig. 6. The figure shows that these CS terms were introduced to MI primarily in the early period. This reveals that the researchers in the MI field begun to apply CS knowledge in the early stage of the field. On the other hand, it also reflects the close relationship between the two fields due to their early contact. It is worth noting that the number of CS terms firstly integrated into MI layers each year is mostly 1 with the maximum of 3. This provides evidence for that the integration of new CS knowledge into MI is a gradual process without an explosive growth. We also observed that the number of new CS terms integrated by all MI layers per year is small, with a maximum of 11 in 1998, as shown in Fig. 7. In general, the number of new CS terms decreased from 1990 to 2016, although with fluctuations in some years. This indicates that less new CS knowledge has been introduced into MI in recent years.

The MI basic layer connected with at most 4 new CS terms one year. Terms from the CS middle layer and the CS detail layer were firstly linked by the layer in 1991. It should be pointed that the two newest terms, "anomaly detecting" and "document analysis", respectively from the CS middle layer and the CS marginal layer were integrated in 2016. The MI middle layer integrated the maximum of 5 CS terms in 1999. It firstly integrated the term, "decision support system" from the CS middle layer in 1990 and the newest term, "document analysis" from the CS marginal layer in 2016. For the MI detail layer, there is only one term that is integrated into the layer from each layer of CS, and a maximum of 3

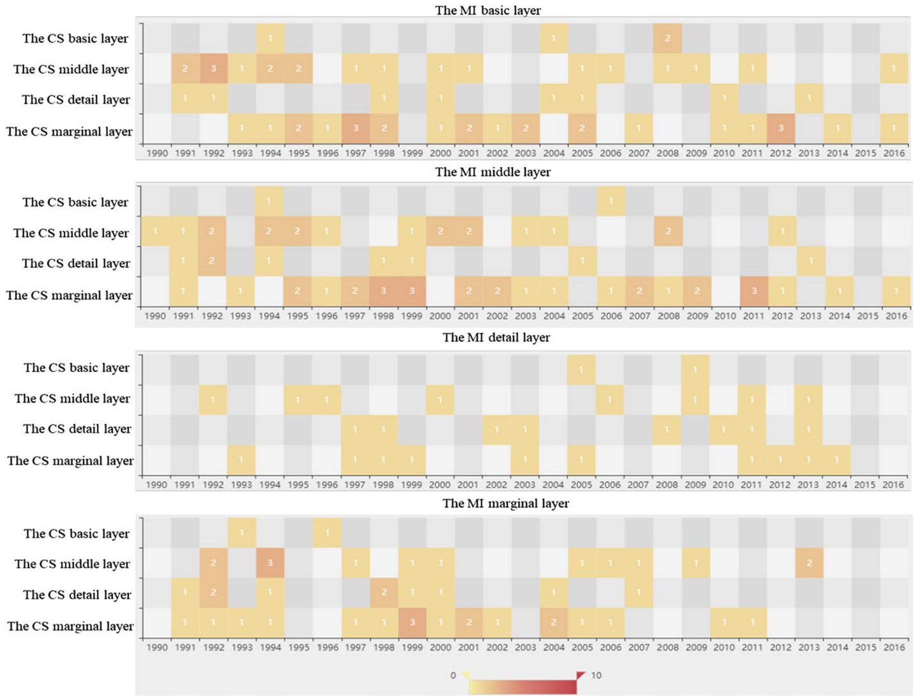


Fig. 6 The time distribution of the CS terms first integrated into the layers of MI from 1987 to 2016

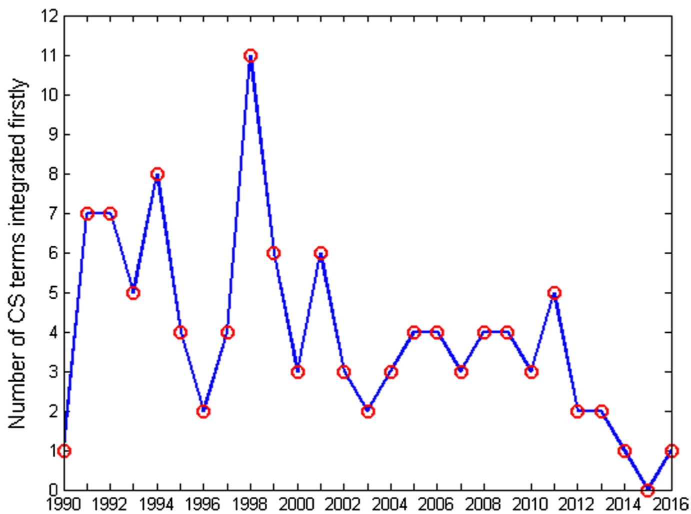


Fig. 7 The number of new CS terms integrated by MI

new CS terms are integrated into the layer in 2011 and 2013. The integration of CS terms into the MI marginal layer is chiefly concentrated in three stages: 1991–1994 (31.11%), 1998–2001 (28.89%) and 2004–2007 (20.00%).

The breadth and strength of knowledge integration

We measured the knowledge integrations of the CS layers by the MI layers at each year in terms of breadth and strength, as presented in Fig. 8. Overall, the breadth and strength of knowledge integrations between MI and CS layers kept increasing from 1990 to 2016, although fluctuations are observed in some years. By comparison, besides the MI detail layer, other MI layers had the broadest and strongest knowledge integration with the CS middle layer over the years, followed by the CS marginal layer. For the MI detail layer, the CS marginal layer showed broader and stronger integration than other layers before 2014. According to the scores, the breadth and strength of knowledge integrations for CS terms in both the MI middle layer and the MI marginal layer were more prominent than in the other two MI layers. In the MI basic layer, the breadth of knowledge integration with the CS middle layer reached a maximum in 2012 and the strongest strength in 2016 (Fig. 8a). However, in other layers of MI, the strength achieved the maximum in 2015. Meanwhile, the breadth in the MI middle layer and the MI detail layer got the maximum value in 2015 as well. This indicates that CS knowledge has been kept merging into the MI layers even in recent years.

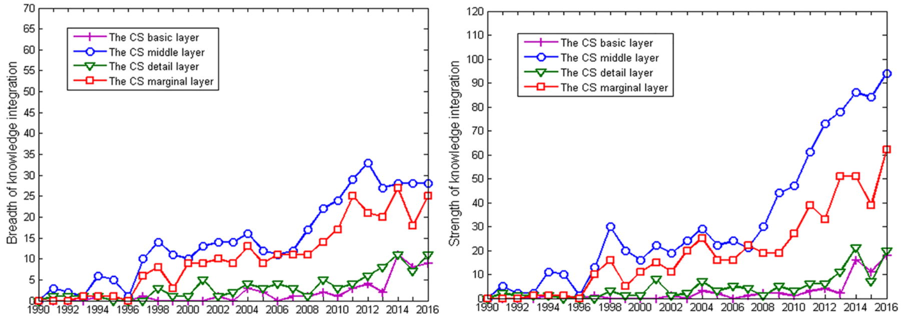
Discussion

Co-word network of a domain provides the benefit of analyzing domain knowledge by a quantitative approach. An important organizing principle of complex network, i.e., hierarchy, was considered in decomposing the knowledge networks of MI and CS in this study. We obtained four layers, including basic, middle, detail and marginal layer, from core to periphery of the networks. A few interesting results were found from the knowledge integrations of CS layers by MI layers. These results provide a detailed and comprehensive understanding of the hierarchical interdisciplinary interactions between MI and CS. The analysis provides a new approach to measuring the knowledge relationships between two fields besides the conventional citation analysis (Karunan et al. 2017).

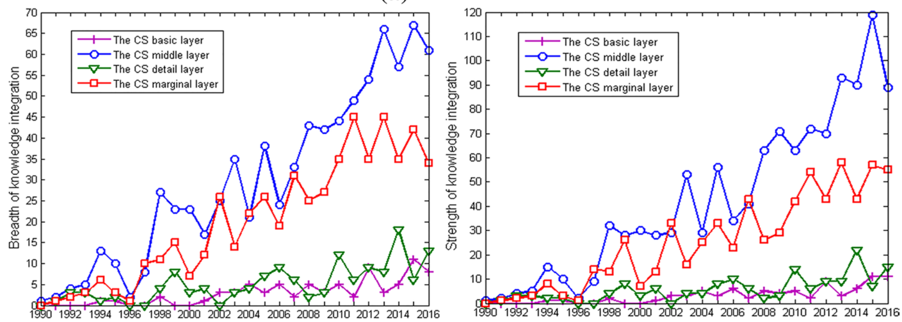
Overall knowledge integration between the layers of MI and CS

The knowledge integration relationships between the layers of MI and CS could be manifested by the shared CS terms. These CS terms represent the knowledge of CS that was integrated by MI. The MI basic layer and the MI middle layer connected more CS terms than the other two layers. The MI basic layer showed the strongest knowledge integration relationship with CS, followed by the middle layer, with the detail layer the weakest. A total of 5 and 35 CS terms have been found in the MI basic and middle layer respectively. These findings evidence that the knowledge from CS has been highly absorbed by the core of MI.

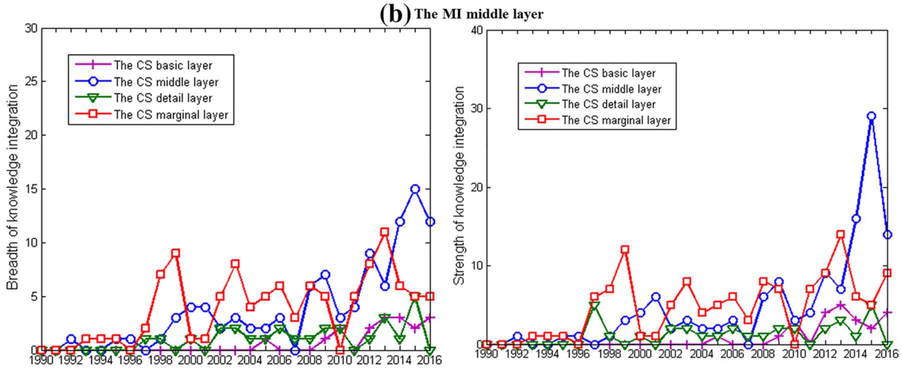
Almost all MI layers have the greatest breadth and strength of knowledge integration with the CS middle layer, followed by the CS marginal layer and detail layer, but with the CS basic layer the weakest. We further examined those trimmed edges with weight of 1 in the knowledge network of MI and obtained the same results. This indicates that the most applied by MI were mediator concepts from CS, rather than basic concepts from CS, which on the contrary was the least. CS basic layer terms are too general concepts in this field,



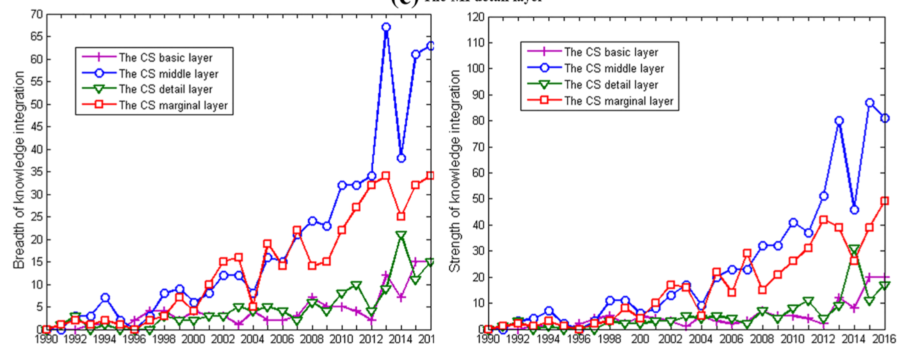
(a) **The MI basic layer**



(b) **The MI middle layer**



(c) **The MI detail layer**



(d) **The MI marginal layer**

Fig. 8 The breadth and strength of knowledge integration between MI and CS layers from 1987 to 2016 such as machine learning and model selection. As an interdisciplinary field, MI borrowed and applied a lot of specific CS techniques and methods to solve the problems in the field. Therefore, these specific CS techniques and methods, e.g., learning algorithm and crowdsourcing, were more manifested as keywords in the research articles. This could explain why the knowledge integrations with the CS detail and marginal layers were broader and stronger than with the CS basic layer. Some CS concept in the middle layer represents a group of techniques, e.g., natural language processing. And some CS middle layer concepts are the shared research problems with MI, such as knowledge management and knowledge representation. These facts result in that the CS middle layer were the closest layer with MI.

Evolution of knowledge integration between the layers of MI and CS

The integration process for CS into MI layers in this paper was examined by investigating the distribution of integrated CS terms and the changing of two metrics over time. The number of CS terms CS integrated by MI assumes an obvious growth tendency. Empirical results suggest that the integration of CS terms by the MI layers over time were in accordance with the overall integration between the layers of MI and CS. Another obvious characteristic is the interdisciplinary interaction between MI and CS became increasingly closer over time, however, fewer new CS terms were integrated by MI in recent years than in earlier years. This implies that recent research in MI utilizes more CS knowledge introduced into MI before to solve the emerging problems in the field, rather than keeps applying new techniques and methods of CS. It's also found that the integration of new CS knowledge into MI is a gradual process without explosive growth, as the distribution of these CS terms first integrated into MI layers are relatively homogeneous. In addition, most of these CS terms in CS knowledge network appeared in early years and were more used in later years, which are in line with their integration by MI over time.

Another important discovery is that the path of knowledge integration between the two fields can be identified by exploring the integration process of CS into MI. The publication year 1998 was a turning point for knowledge integration across layers of MI and CS. Terms integrated before 1998 were mainly from the CS middle layer, and they were mostly tied with the MI basic and middle layers. From 1998 on, a larger number of terms from the CS marginal layer were integrated by MI layers, except for the MI detail layer. It reveals that when applied in MI, higher level knowledge of CS was ahead of lower level knowledge in the time series. This analysis provides a new perspective for identifying knowledge evolutionary path based on the distribution of integrated CS terms and breadth and strength of their integration, rather than the traditional bibliographic citation information (Lucio-Arias and Leydesdorff 2008) or topic relevance analysis (Wang et al. 2014).

Conclusion

In this paper, we put forward a new method to explore the knowledge integration of CS by MI and its evolution from the perspective of hierarchy. The hierarchical structure of the MI knowledge network was examined by disclosing basic properties and shared CS terms in the four decomposed layers. The interdisciplinary interactions between different layers of MI and CS were analyzed by observing the connected CS terms and quantifying the breadth and strength of knowledge integration. The evolutionary characteristics on the knowledge integration of CS by MI over time are also revealed. Especially, the knowledge integration path is analyzed to demonstrate the process that the layers of CS knowledge were integrated by MI. The findings of this study provide a clearer and much deeper understanding about the integration of CS knowledge into the interdisciplinary field, MI, from a hierarchical perspective. The methodology and metrics used in this study could be applied to examine knowledge integration relationship between pairwise fields.

Admittedly, some limitations could be identified in this study. First, although showing a high consistency for the discipline categorization of terms among the three annotators, we manually annotated the terms belonging to computer science, which still suffers from subjectivity. Second, the knowledge in one discipline often is from multiple disciplines. In this study, we did not differentiate the terms in MI except the CS terms. Some of them could come from medicine, biology, or other disciplines. We did not explore how CS terms combine with the terms of other disciplines yet. In fact, this limitation provides clues for our future study. We will explore knowledge interactions among MI and other disciplines based on the proposed methods to further understand how multiple disciplines take different roles and contribute to the interdisciplinary field of MI. What's more, we plan to distinguish the functions of keywords (Lu et al. 2018) to discover the patterns of knowledge integration among multiple fields in an interdisciplinary field.

Acknowledgements This study was supported by the National Natural Science Foundation of China (71790612 and 71804135). We gratefully thank Dr. Zhe He in the School of Information at the Florida State University to help label the discipline information of the selected terms.

Appendix A

See Table 8.

Table 8 Journal list in "Medical Informatics" subject category in WoS

No.	Journal name	No.	Journal name
1	Applied Clinical informatics	13	International Journal of Technology Assessment in Health Care
2	Artificial Intelligence in Medicine	14	Journal of Biomedical Informatics
3	Biomedical Engineering-Biomedizinische Technik	15	Journal of Evaluation in Clinical Practice
4	BMC Medical Informatics and Decision Making	16	Journal of Medical Internet Research
5	Cin-Computers Informatics Nursing	17	Journal of Medical Systems
6	Computer Methods and Programs in Biomedicine	18	Journal of The American Medical Informatics Association
7	Health Informatics Journal	19	Medical & Biological Engineering & Computing
8	Health Information Management Journal	20	Medical Decision Making
9	IEEE Journal of Biomedical and Health Informatics	21	Methods of Information in Medicine
10	IEEE Transactions on Information Technology in Biomedicine	22	Statistical Methods in Medical Research
11	Informatics for Health & Social Care	23	Statistics in Medicine
12	International Journal of Medical Informatics	24	Therapeutic Innovation & Regulatory Science

Appendix B

See Table 9.

Table 9 The list of 30 Computer Science journals

No.	Journal name	No.	Journal name
1	Artificial Intelligence In Medicine	16	International Journal of Computer Vision
2	IEEE Transactions on Pattern Analysis and Machine Intelligence	17	Image and Vision Computing
3	Expert Systems with Applications	18	Data Mining and Knowledge Discovery
4	Machine Learning	19	Computational Linguistics
5	Artificial Intelligence	20	Data & Knowledge Engineering
6	Pattern Recognition	21	Statistical Analysis and Data Mining
7	Medical Image Analysis	22	Soft Computing
8	IEEE Transactions on Image Processing	23	Siam Journal on Imaging Sciences
9	Pattern Recognition Letters	24	Semantic Web
10	Journal of Machine Learning Research	25	Pattern Analysis and Applications
11	IEEE Transactions on Knowledge and Data Engineering	26	Neural Processing Letters
12	Neural Computation	27	Neural Network World
13	Neural Networks	28	Neural Computing & Applications
14	Decision Support Systems	29	Network-Computation in Neural Systems
15	Neurocomputing	30	Wiley Interdisciplinary Reviews-Data Mining and Knowledge Discovery

Appendix C

See Table 10.

Table 10 The distribution of 105 CS terms in CS network and their integration by MI

No.	Terms	Layer	Breadth	Strength	No.	Terms	Layer	Breadth	Strength
1	Decision making	Middle	46	140	54	Big data	Middle	2	4
2	Decision support system	Middle	27	80	55	Data security	Marginal	2	4
3	Natural language processing	Middle	21	61	56	Crowd-sourcing	Marginal	2	4
4	Information system	Middle	20	50	57	Problem based learning	Marginal	2	5
5	Social networking sites	Middle	20	52	58	Learning algorithm	Detail	2	9
6	Agility	Middle	17	42	59	Security	Middle	2	11
7	World wide web	Marginal	16	38	60	Bar code	Marginal	2	6
8	DEA analysis	Marginal	14	43	61	Software design	Marginal	2	4
9	Web site	Detail	11	32	62	Data presentation	Marginal	2	4
10	Online system	Detail	11	23	63	LCModel	Marginal	2	4
11	User computer interface	Marginal	11	24	64	Group recommender system	Marginal	2	4
12	Data quality	Marginal	10	23	65	Machine learning	Basic	2	5
13	Document analysis	Marginal	10	30	66	Data banks	Marginal	2	4
14	Informatics technology	Middle	9	20	67	Factor analysis	Detail	2	4
15	Non linear dynamics	Detail	8	23	68	Social network analysis	Detail	2	4
16	Model selection	Basic	8	35	69	Activity recognition	Marginal	2	5
17	Text mining	Middle	7	16	70	Inverse probability weighting	Middle	2	8
18	Sensitivity analyses	Detail	7	17	71	Interpolation artefact	Middle	1	4
19	Computer security	Marginal	7	14	72	Online intervention	Marginal	1	2
20	Knowledge management	Middle	7	23	73	Partitioning	Marginal	1	2
21	Knowledge representation	Middle	6	21	74	Support vector machine	Basic	1	2
22	Information retrieval	Middle	6	29	75	Missing data	Marginal	1	2
23	Knowledge acquisition	Detail	5	10	76	Structured data	Marginal	1	2
24	Motion learning	Marginal	5	11	77	Bench marking	Marginal	1	2
25	Facial feature detection	Marginal	5	21	78	Artificial intelligence	Middle	1	2
26	Spiking neuron	Detail	5	21	79	Computer assisted decision making	Middle	1	2
27	Distance education	Detail	5	12	80	Web-based services	Marginal	1	2

Table 10 (continued)

No.	Terms	Layer	Breadth	Strength	No.	Terms	Layer	Breadth	Strength
28	Forensic identification	Marginal	5	20	81	Expert system	Middle	1	2
29	Human computer interaction	Detail	5	10	82	Mixture model	Basic	1	2
30	Access control	Marginal	5	10	83	Knowledge discovery	Middle	1	3
31	FIR filter	Marginal	5	19	84	Salience	Middle	1	4
32	Semantic net	Middle	4	8	85	Knowledge translation	Marginal	1	2
33	Data visualization	Middle	4	8	86	Multi-spectral imaging	Marginal	1	2
34	Classification	Basic	4	9	87	Neural network	Basic	1	3
35	R software	Marginal	4	8	88	Random effect model	Marginal	1	2
36	Multi-media systems	Detail	4	9	89	Fully automated	Marginal	1	2
37	Adaptive design	Marginal	4	9	90	Word sense disambiguation	Middle	1	2
38	Coreference	Marginal	4	12	91	Automated detection	Marginal	1	2
39	Bayesian analyses	Middle	4	8	92	Bi-clustering	Basic	0	0
40	Semantic interoperability	Middle	4	10	93	Feature detection	Detail	0	0
41	Computer simulation	Marginal	3	7	94	Functional imaging	Detail	0	0
42	Complex network	Detail	3	6	95	Gibbs sampling	Detail	0	0
43	Information extraction	Marginal	3	6	96	Hadoop	Marginal	0	0
44	Usability testing	Middle	3	7	97	Knowledge base	Marginal	0	0
45	pRAM nets	Marginal	3	17	98	Maximum likelihood estimation	Marginal	0	0
46	Anomaly detecting	Mddle	3	9	99	Model predictive control	Marginal	0	0
47	Cloud computing	Marginal	3	9	100	Pattern analysis	Basic	0	0
48	System modeling	Detail	3	11	101	Semantic similarity	Marginal	0	0
49	Data coding	Marginal	3	7	102	Virtual reality	Marginal	0	0
50	Administrative database	Marginal	3	6	103	Web service	Detail	0	0
51	Decision analyses	Marginal	3	8	104	Active contour	Middle	0	0
52	System development	Marginal	3	6	105	Conditional random fields	Marginal	0	0
53	Complex adaptive Systems	Marginal	3	7					

References

Abramo, G., D'Angelo, C. A., & Costa, F. D. (2012). Identifying interdisciplinarity through the disciplinary classification of coauthors of scientific publications. *Journal of the Association for Information Science & Technology, 63*(11), 2206–2222.

Alvarez-Hamelin, J. I., Dall'Asta, L., Barrat, A., & Vespignani, A. (2005). k-core decomposition: a tool for the visualization of large scale networks. arXiv:cs/0504107.

Alvarez-Hamelin, J. I., Dall'Asta, L., Barrat, A., & Vespignani, A. (2017). K-core decomposition of Internet graphs: hierarchies, self-similarity and measurement biases. *Networks & Heterogeneous Media, 3*(2), 371–393.

Bergmann, T., Dale, R., Sattari, N., Heit, E., & Bhat, H. S. (2016). The interdisciplinarity of collaborations in cognitive science. *Cognitive Science, 41*(5), 1412–1418.

Besselaar, P. V. D., & Heimeriks, G. (2006). Mapping research topics using word-reference co-occurrences: A method and an exploratory case study. *Scientometrics, 68*(3), 377–393.

Braun, T., & Schubert, A. (2003). A quantitative view on the coming of age of interdisciplinarity in the sciences 1980–1999. *Scientometrics, 58*(1), 183–189.

Carmi, S., Havlin, S., Kirkpatrick, S., Shavitt, Y., & Shir, E. (2007). A model of Internet topology using k-shell decomposition. *Proceedings of the National Academy of Sciences of the United States of America, 104*(27), 11150–11154.

Chi, R., & Young, J. (2013). The interdisciplinary structure of research on intercultural relations: a co-citation network analysis study. *Scientometrics, 96*(1), 147–171.

Clauset, A., Moore, C., & Newman, M. E. (2008). Hierarchical structure and the prediction of missing links in networks. *Nature, 453*(7191), 98–101.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46.

Cole, S. (1983). The hierarchy of the sciences? *American Journal of Sociology, 89*(1), 111–139.

Collins, J. J., & Chow, C. C. (1998). It's a small world. *Nature, 393*(6684), 409–410.

Dong, K., Xu, H., Rui, L., Ling, W., & Shu, F. (2018). An integrated method for interdisciplinary topic identification and prediction: a case study on information science and library science. *Scientometrics, 115*(2), 849–868.

Dorogovtsev, S. N., Goltsev, A. V., & Mendes, J. F. F. (2006). k-core organization of complex networks. *Physical Review Letters, 96*(4), 040601.

Eidsaa, M., & Almaas, E. (2013). S-core network decomposition: A generalization of k-core analysis to weighted networks. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics, 88*(6), 062819.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy Sciences of the United States of America, 101*(1), 5228–5235.

Han, J. W., Pei, J., Yin, Y., Mao, R., et al. (2004). Mining frequent patterns without candidate generation: A frequent-pattern tree approach. *Data Mining and Knowledge Discovery, 8*(1), 53–87.

Hirst, G. (1978). Discipline impact factors: A method for determining core journal lists. *Journal of the American Society for Information Science and Technology, 29*(4), 171–172.

Hu, J., & Zhang, Y. (2017). Discovering the interdisciplinary nature of big data research through social network analysis and visualization. *Scientometrics, 112*(1), 91–109.

Karlovčec, M., & Mladenić, D. (2015). Interdisciplinarity of scientific fields and its evolution based on graph of project collaboration and co-authoring. *Scientometrics, 102*(1), 433–454.

Karunan, K., Lathabai, H. H., & Prabhakaran, T. (2017). Discovering interdisciplinary interactions between two research fields using citation networks. *Scientometrics, 113*(1), 335–367.

Khaouid, W., Barsky, M., Srinivasan, V., & Thomo, A. (2015). K-core decomposition of large networks on a single PC. *Proceedings of the Vldb Endowment, 9*(1), 13–23.

Lancichinetti, A., & Fortunato, S. (2012). Consensus clustering in complex networks. *Scientific Reports, 2*(13), 336.

Lee, B., & Jeong, Y. I. (2008). Mapping Korea's national R&D domain of robot technology by using the co-word analysis. *Scientometrics, 77*(1), 3–19.

Lee, K., Jung, H., & Song, M. (2016). Subject–method topic network analysis in communication studies. *Scientometrics, 109*(3), 1–27.

Lee, P. C., Su, H. N., & Chan, T. Y. (2010). Assessment of ontology-based knowledge network formation by vector-space model. *Scientometrics, 85*(3), 689–703.

Leeuwen, T. N. V., & Moed, H. F. (2005). Characteristics of journal impact factors: The effects of uncitedness and citation distribution on the understanding of journal impact factors. *Scientometrics, 63*(2), 357–371.

Leydesdorff, L., & Rafols, I. (2011). Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. *Journal of Informetrics, 5*(1), 87–100.
Li, L. G., & Zhang, M. C. (2017). The construction of top disciplines should attach importance to developing discipline group and cross-discipline. Retrieved from http://epaper.gmw.cn/gmrb/html/2017-09/04/nw.D110000gmrb_20170904_1-11.htm. (**in Chinese**).
Liu, G. Y., Hu, J. M., & Wang, H. L. (2012). A co-word analysis of digital library field in China. *Scientometrics*, *91*(1), 203–217.
Liu, Y., Tang, M., Zhou, T., & Do, Y. (2015). Core-like groups result in invalidation of identifying super-spreader by k-shell decomposition. *Scientific Reports, 5,* 9602.
Liu, Z., & Wang, C. (2005). Mapping interdisciplinarity in demography: a journal network analysis. *Journal of Information Science, 31*(4), 308–316.
Liu, P., & Xia, H. (2015). Structure and evolution of co-authorship network in an interdisciplinary research field. *Scientometrics, 103*(1), 101–134.
Lu, W., Huang, Y., Bu, Y., & Cheng, Q. (2018). Functional structure identification of scientific documents in computer science. *Scientometrics, 115*(3), 1–24.
Lucio-Arias, D., & Leydesdorff, L. (2008). Main-path analysis and path-dependent transitions in HistCite-based historiograms. *Journal of the Association for Information Science & Technology, 59*(12), 1948–1962.
Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., & Mcclosky, D. (2014). The Stanford CoreNLP Natural Language Processing Toolkit. *Proceedings of 52Nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. https://doi.org/10.3115/v1/P14-5010.
Mao, J., Lu, K., Zhao, W., & Cao, Y. (2018). How many keywords do authors assign to research articles–a multi-disciplinary analysis?. In *Conference 2018 Proceedings*.
Morillo, F., Bordons, M., & Gómez, I. (2001). An approach to interdisciplinarity through bibliometric indicators. *Scientometrics, 51*(1), 203–222.
Morris, T. A., & Mccain, K. W. (1998). The structure of medical informatics journal literature. *Journal of American Medical Informatics Association, 5*(5), 448–466.
Nanni, L., Paci, M., Brahnam, S., & Ghidoni, S. (2017). An ensemble of visual features for Gaussians of local descriptors and non-binary coding for texture descriptors. *Expert Systems with Applications*, *82*(C), 27–39.
Nichols, L. G. (2014). A topic model approach to measuring interdisciplinarity at the National Science Foundation. *Scientometrics, 100*(3), 741–754.
Piepenbrink, A., & Nurmammadov, E. (2015). Topics in the literature of transition economies and emerging markets. *Scientometrics, 102*(3), 2107–2130.
Porter, A. L., & Chubin, D. E. (1985). An indicator of cross-disciplinary research. *Scientometrics, 8*(3–4), 161–176.
Porter, A. L., Cohen, A. S., Roessner, J. D., & Perreault, M. (2007). Measuring researcher interdisciplinarity. *Scientometrics, 72*(1), 117–147.
Rafols, I., & Meyer, M. (2010). Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience. *Scientometrics, 82*(2), 263–287.
Sales-Pardo, M., Guimerà, R., Moreira, A. A., & Amaral, L. A. (2007). Extracting the hierarchical organization of complex systems. *Proceedings of the National Academy of Sciences of the United States of America, 104*(39), 15224–15229.
Salton, G., & Mcgill, M. J. (1988). The SMART and SIRE experimental retrieval systems. In *Readings in information retrieval*, 1988.
Schummer, J. (2004). Multidisciplinarity, interdisciplinarity, and patterns of research collaboration in nanoscience and nanotechnology. *Scientometrics, 59*(3), 425–465.
Song, M., & Kim, S. Y. (2013). Detecting the knowledge structure of bioinformatics by mining full-text collections. *Scientometrics*, *96*(1), 183–201.
Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society, Interface, 4*(15), 707.
Valentin, F., Norn, M. T., & Alkaersig, L. (2016). Orientations and outcome of interdisciplinary research: the case of research behavior in translational medical science. *Scientometrics*, *106*(1), 1–24.
Wang, X., Cheng, Q., & Lu, W. (2014). Analyzing evolution of research topics with NEViewer: A new method based on dynamic co-word networks. *Scientometrics, 101*(2), 1253–1271.
Wang, L., Notten, A., & Surpatean, A. (2013). Interdisciplinarity of nano research fields: A keyword mining approach. *Scientometrics, 94*(3), 877–892.
White, H. D., & McCain, K. W. (1997). Visualization of Literatures. *Annual Review of Information Science and Technology, 32*(1), 99–168.

Xiao, L., Chen, G., Sun, J., Han, S., & Zhang, C. (2016). Exploring the topic hierarchy of digital library research in China using keyword networks: A K-core decomposition approach. *Scientometrics, 108*(3), 1–17.
Xu, H., Guo, T., Yue, Z., Ru, L., & Fang, S. (2016). Interdisciplinary topics of information science: a study based on the terms interdisciplinarity index series. *Scientometrics, 106*(2), 583–601.
Zhang, H., Zhao, H., Cai, W., Liu, J., & Zhou, W. (2010). Using the k-core decomposition to analyze the static structure of large-scale software systems. *Journal of Supercomputing, 53*(2), 352–369.