Research Paper

Identification and Prediction of Interdisciplinary Research Topics: A Study Based on the Concept Lattice Theory

Haiyun Xu[1,2†], Chao Wang[3,4], Kun Dong[5], Zenghui Yue[6]

[1]Institute of Scientific and Technical Information of China, Beijing 100038, China
[2]Chengdu Documentation and Information Center, Chinese Academy of Sciences, Chengdu 610041, China
[3]Information research institute of Shandong Academy of sciences, Jinan 250014, China
[4]Qilu University of Technology, Jinan 250353, China
[5]Science and Technology Information Research Institute, Shandong University of Technology, Zibo 255091, China
[6]School of Medical Information Engineering, Jining Medical University, Rizhao 276826, China

Abstract

Purpose: Formal concept analysis (FCA) and concept lattice theory (CLT) are introduced for constructing a network of IDR topics and for evaluating their effectiveness for knowledge structure exploration.

Design/methodology/approach: We introduced the theory and applications of FCA and CLT, and then proposed a method for interdisciplinary knowledge discovery based on CLT. As an example of empirical analysis, interdisciplinary research (IDR) topics in Information & Library Science (LIS) and Medical Informatics, and in LIS and Geography-Physical, were utilized as empirical fields. Subsequently, we carried out a comparative analysis with two other IDR topic recognition methods.

Findings: The CLT approach is suitable for IDR topic identification and predictions.

Research limitations: IDR topic recognition based on the CLT is not sensitive to the interdisciplinarity of topic terms, since the data can only reflect whether there is a relationship between the discipline and the topic terms. Moreover, the CLT cannot clearly represent a large amounts of concepts.

Practical implications: A deeper understanding of the IDR topics was obtained as the structural and hierarchical relationships between them were identified, which can help to get more precise identification and prediction to IDR topics.

Originality/value: IDR topics identification based on CLT have performed well and this theory has several advantages for identifying and predicting IDR topics. First, in a concept lattice, there is a partial order relation between interconnected nodes, and consequently, a complete concept

---

† Corresponding author: Haiyun Xu (E-mail: xuhy@clas.ac.cn).

lattice can present hierarchical properties. Second, clustering analysis of IDR topics based on concept lattices can yield clusters that highlight the essential knowledge features and help display the semantic relationship between different IDR topics. Furthermore, the Hasse diagram automatically displays all the IDR topics associated with the different disciplines, thus forming clusters of specific concepts and visually retaining and presenting the associations of IDR topics through multiple inheritance relationships between the concepts.

1   Introduction

With the development of science and technology, scientific research is no longer limited to the study of a single field but is extended to interdisciplinary research (IDR). Disciplines show a trend whereby they are both highly differentiated from and integrated with each other. IDR is a comprehensive scientific activity promoted by the developmental needs of both our society and the discipline itself (Klein, 2000). In fact, IDR topic identification can help analyze new technological research frontiers and hotspots, and predict the growth of novel disciplines. IDR topic trend forecasting can help understand the absorption and diffusion trends of knowledge from different disciplines, and can thereby enhance interdisciplinary collaboration and promote the integration and development of disciplines.

Scientific literatures are an important data source for IDR. It is increasingly important to develop effective knowledge-mining methods for identifying IDR topics from large volumes of scientific literatures and to predict the future trends. The existing studies identify IDR topics mainly using co-occurrence networks, which cannot represent the structural and hierarchical relationships between IDR topics. If we can clarify the features of the hierarchical structure between IDR topics, a deeper understanding of the IDR topics can be obtained, and the prediction of IDR topics will be more precise.

FCA is a data analysis method that can reduce the data complexity while retaining nearly all the details of the data (Ganter & Wille, 1997). FCA is used to construct a concept lattice and reconstruct a bipartite network as a hierarchical network. A concept lattice based on domain knowledge can represent the knowledge structure in specific areas, thereby showing the hierarchy and relationships between knowledge units, which allows the discovery of implicit knowledge. Nowadays, the construction of a concept lattice for domain knowledge is used for knowledge representation; however, FCA and concept lattice are not widely used for the identification of IDR topics. In this study, FCA and concept lattice theory (CLT) are introduced for constructing a network of IDR topics and for evaluating their effectiveness for knowledge structure exploration.

This paper is organized as follows. First, various methods and existing problems of IDR topic identification research are reviewed, and the characteristics of different methods used for IDR topic analysis are clarified. Then, the theory and application of FCA and CLT are introduced, and the method for identifying and predicting IDR topics based on the CLT is described. Subsequently, IDR topics in LIS & Medical Informatics, and in LIS & Geography-Physical, were utilized as empirical fields. A comparative analysis with two other IDR topic recognition methods is performed. Finally, the advantages and limitations of this study are discussed.

2   The art of interdisciplinary research

2.1   Bibliometrics methods

The quantitative research in Scientometrics on the IDR is mainly by means of bibliometrics. The current quantitative research on the IDR mainly depends on the characteristics of one or several disciplines, journals, or researchers, and where research papers and references are mostly the objects of analysis (Porter et al., 2008; Schummer, 2004). Small (1973) noted that specific areas of a discipline could be analyzed by a co-citation network of literatures and co-citation analysis, which has provided an effective way of thinking about interdisciplinary cross-referencing. Some scholars have also studied interdisciplinarity among different research publications by assessing the cooperation of authors from different disciplines (Qiu, 1992). Hammarfelt (2011) made a statistical analysis of citation topics from various literary journals by citation analysis, compared changing trends in two periods, and measured interdisciplinarity in a discipline by citation topics.

2.2   Visualization methods

Visualization is an important way of in IDR topic identification. On the basis of the intersection of standardized topic keywords, Min and Sun (2014) drew a dendrogram and strategic diagram of cross keywords; they discussed the internal relations and development context of interdisciplinary research hotpots using cluster analysis and strategy analysis introduced from quantitative angles, and combined these with an adhesion index named by clustering class groups. Based on the problem of overlapping community recognition in the interdisciplinary fields, Li et al. (2013) mined the cross-research topics between Library and Information Science (LIS) and computer science using a complex network discovery tool named CFinder, which made an analysis of the cross-research topics by a visual display of clustering and overlapping social networks between these two disciplines. Zhang et al. (2011) undertook empirical research on interdisciplinarity by analyzing ten years of literatures from the domestic core journals of LIS and computer science, constructed the network of both authors and literatures based on citation relationships,

and further discussed the cross-disciplinary relationships between the research
studies in these two disciplines.

2.3 IDR measurement index

The IDR index is the important tool to mining IDR topics. Chang and Huang
(2012) analyzed the interdisciplinarity of LIS by the Brillouin index, proved its
validity for measuring interdisciplinarity in LIS, and drew the conclusion that the
interdisciplinary level of LIS has strengthened. Leydesdorff and Rafols (2011)
utilized the Gini coefficient, information entropy, and Rao-Stirling indicators to
measure the interdisciplinary features of academic journals. Leydesdorff et al.
(2013) visually displayed both the citing and cited interdisciplinary citation matrices
using Rao-Stirling indicators. Xu et al. (2016) introduced a measurement index
called topic terms interdisciplinarity (TI) for IDR topic mining. They showed that
the TI value can identify IDR topic terms effectively. The integrated utilization of
a variety of indicators is important for mining the IDR topics.

To summarize, the studies mentioned above mainly identified IDR topics using
co-occurrence networks, which are not good at representing the structural and
hierarchical relationships between IDR topics. If the structural and hierarchical
relationships between IDR topics can be identified, a deeper understanding of the
IDR topics can be obtained, and the prediction of IDR topics could acquire more
supported information.

FCA is one of data analysis methods, which does not artificially reduce the data
complexity, and thus, it retains nearly all of the details of the data (Ganter & Wille,
1997). FCA is used to construct a concept lattice and reconstruct a bipartite network
as a hierarchical network. FCA and concept lattice theory (CLT) are both powerful
tools for conceptualization during knowledge processing, and in addition, the
construction of a concept lattice for domain knowledge is a method for knowledge
representation. However, FCA and CLT are not widely used for the identification
of IDR topics.

3 Research and application of FCA and the concept lattice

3.1 Mathematical foundations of FCA and the concept lattice

FCA was first proposed at the Technical University of Darmstadt in 1982 by
Wille, who also reconstructed the lattice theory with FCA (Wille, 2009). From
a philosophical perspective, FCA and the concept lattice are extensions and
implementations of the philosophy of Ludwig Wittgenstein, who is the founder of
the philosophy of logic and the philosophy of language. Wille stated that concepts
are the basic units of human thoughts, so the structures of the logic and information

Research Paper

are both based on the concepts and conceptual systems (Wille, 2002). FCA is a data analysis method through extracting a hierarchical structure of clusters from tabular data describing objects with their attributes, foundations, algorithms, and a survey of applications. Also FCA is a method for visualizing the internal structures and relationships of data (Carpineto et al., 2005; Teng, 2012; Wille 2009). The mathematical variables and relationships can be described by four definitions, as follows.

Data analysis using FCA always starts with a formal context defined as a triple (G, M, I), where G is a set of formal objects, M is a set of formal attributes, and I is the relation between G and M (i.e., $I \subseteq G \times M$). A formal context is generally represented as a binary incidence table, where the crosses represent the binary relation between the object set and the attribute set. For the formal context, operators $\uparrow: 2^G \rightarrow 2^M$ and $\downarrow: 2^M \rightarrow 2^G$ are defined for every $A \subseteq G$ and $B \subseteq M$ by $A\uparrow = \{m \in M / \text{for each } g \in A: <g, m> \in I\}$, $B\downarrow = \{g \in G / \text{for each } m \in B: <g, m> \in I\}$. The operators $\uparrow$ and $\downarrow$ are known as concept forming operators. A formal concept of a formal context is defined as an ordered pair (A, B) with $A \subseteq G$ and $B \subseteq M$ such that $A\uparrow = B$ and $B\downarrow = A$. We refer to A and B as an extent and intent, respectively, of the formal concept (A, B). A formal concept (A, B) of the context (G, M, I) is defined as a subconcept of the formal concept (C, D) of (G, M, I) and (C, D) as a superconcept of (A, B) if the extent A is contained in the extent C and, equivalently, if the intent B contains the intent D. The set of all concepts of a context (G, M, I) with the order relation $\leq$ is always a complete lattice called the concept lattice.

FCA organizes the information through concept lattice, which fundamentally comprises a partial order modeling the subconcept–superconcept hierarchy (Jamil & Deogun, 2001). CLT holds that the real world is comprised of various concepts and that a concept usually contains a connotation and extension. The connotation of a concept embodies the characteristic attributes of cognitive things, and an extension embodies the objects of cognitive things. As a mathematical abstract of concept systems, a concept lattice can help find information and thus create knowledge (Wille, 2002). Graphically represented concept lattices have proved to be useful in discovering and understanding the conceptual relationships in given data sets.

3.2 The theory and applications of concept lattice in knowledge discovery

The concept lattice for knowledge discovery in databases was first used by Venter (Venter, Oosthuizen, & Roos, 1997), who described the visualization and a guided role for the concept lattice during the process of knowledge discovery. FCA aims to find concept clusters in data sets of formal concepts, using which the data attribute associations can be found and displayed as a concept lattice (Belohlavek & Vychodil,

2009). The concept lattice of domain knowledge utilizes the partial order of objects and the attributes of the concepts to represent knowledge nodes in a hierarchical structure. In a concept lattice, the extension and connotation of a concept can be discovered from the hyponymy semantic relation between nodes, and the knowledge structure can be organized through a hierarchy and network-oriented approach. Moreover, according to the partial order relation of a concept lattice, the top node has the broadest knowledge connotation and extension in the concept lattice and can be regarded as the core of the research domain. In contrast to other relationships of partial order, such as a tree of knowledge organization system, multiple inheritance features exist in a concept lattice. Therefore, the concept lattice structure can support knowledge discovery in databases (KDD) (Stumme, 2009). Kumar (2011) conducted KDD based on random projections by the FCA, and confirmed the effectiveness of this method based on an empirical analysis of medical data sets. Cimiano used a concept lattice to automatically classify a text corpus and obtain hierarchy concepts (Cimiano, Hotho, & Staab, 2005).

Previous knowledge discovery studies based on FCA generally aimed to construct related keywords in the form of coupling networks, such as an author-keywords coupling network. These studies first determined the formal concept background for the research topics and thereafter established a concept lattice to discover knowledge. Here we list the representative studies: Liu et al. proposed keywords-author coupling analysis based on FCA and their empirical analysis showed that compared with the traditional co-word analysis method, author-keywords coupling analysis method based on FCA could obtain better hierarchical effects and reveal fine-grained knowledge structures (Liu & Wang, 2012; Liu & Wu, 2014). Gao et al. used a concept lattice based on a literature-coupling network, visually represented the knowledge structure and component heterogeneity of the coupling network (Gao, 2012). Teng et al. (2011) proposed granularity concept analysis (GCA) based on FCA for keyword analysis and built different granularity concept lattices with GCA. By mining different granularity concept lattices, they analyzed high frequency keywords and relatively low-frequency keywords in the ontology field, and analyzed the structure of knowledge and inherent associations (Teng, Bi, & Bao, 2011).

3.3 Visualization of a concept lattice

Concept lattice analysis has its own visualization tool–Hasse diagram (Ganter & Wille, 2012). In a Hasse diagram, each node represents one concept and there is a partial order relationship between the nodes, so the visualization of partial order relationships can facilitate the hierarchical and graphical display of knowledge topics. This visualization method describes the relationship between the concepts

and topics based on human cognitive law, thereby providing a graded visual effect that is easier to understood and explained.

Drawing a Hasse diagram manually requires many tasks, however a variety of automatic generation tools are available now. Teng et al. listed the main tools for concept lattice building, including ConExp, Lattice Miner, Lattice Navigator, ToscanaJ, and Coron (Teng et al., 2012). A comparative analysis of ConExp and LatticeMiner was also performed (Lahcen & Kwuida, 2010), which are the two most commonly used concept lattice-building tools (Teng & Bi, 2010). Shao et al. proposed a method of interdisciplinary knowledge structure detection based on concept lattice and bibliographic coupling. They determined the correlated characteristics of each research topic and core corresponding author associated with each topic, using association rule mining and hierarchical clustering based on a concept lattice (Shao & Li, 2015).

4   Theory of IDR topics based on CLT

After the above characteristic analysis of the CLT, especially the advantages of KDD supported by CLT, we further propose two major merits in the identification and prediction of IDR topics based on the CLT, With the main reasons as follows.

First, a concept lattice can provide a good representation mode to a knowledge unit of IDR topics. The formal conceptual background describes the relationships between knowledge units. A concept lattice based on domain knowledge represents the knowledge structure in specific areas, thereby showing the hierarchy and relationships between knowledge units of IDR topics, which allows the discovery of implicit knowledge of IDR topics.

In a concept lattice, there is a partial order relation between interconnected nodes, and consequently, a complete concept lattice can present hierarchical properties. When we regard topics as objects and disciplines as attributes and build a concept lattice, the concepts of knowledge structure comprise disciplines (extension) and topics (connotation). The multiple inheritance relationships between hierarchical concepts can reveal the potential complex relationships between two or more disciplines. Therefore, a concept lattice can reflect the clusterings and characteristics associated with the concept, which can visualize the disciplinary knowledge structure and identify the IDR topics. Moreover, through the evolution trajectory of IDR topics, the future IDR topics can be predicted.

Second, clustering analysis of IDR topics based on concept lattice can yield clusters that highlight the essential knowledge features and help display the semantic relationship between different IDR topics.

The cluster analysis based on concept lattice involves recognizing interesting formal concepts. FCA aims at discovering conceptual clusters and visualizing them

based on a conceptual structure called the concept lattice (Belohlavek & Vychodil, 2009). IDR topic recognition based on a concept lattice is usually with the formal conceptual background of a discipline-topic terms binary co-occurrence network which is essentially a dichotomous network analysis, while the CLT has its own advantages for IDR topic identification for it makes more detailed information available for the analysis. To be specific, a concept is presented in the form of a starting node that is extended, spread, and eventually converged to a node. The network extension process is the intersection process between the discipline and the topic terms. Therefore, the hierarchical presentation of a concept lattice expands the co-occurrence network analysis, which can help display the semantic distance and relationship between different IDR topics.

Furthermore, the Hasse diagram automatically displays all the IDR topics associated with the different disciplines, thus forming clusters of specific concepts and visually retaining and presenting the associations of IDR topics through multiple inheritance relationships between the concepts. We can analyze any IDR topic that intersects with its specific crossed disciplines in a hierarchical network.

5   Method of IDR topic discovery based on CLT

5.1   Main steps of IDR topic discovery based on CLT

The overall process of IDR topic discovery based on CLT can be roughly divided into two parts: knowledge units representation based on a concept lattice and topic analysis based on cluster analysis. In this study, we use the discipline–topic terms in a formal context, conduct cluster analysis of the concept lattice to reflect the structure and hierarchy of the knowledge network, and consider the method employed for constructing a network of IDR topics and evaluate its effectiveness for knowledge structure exploration.

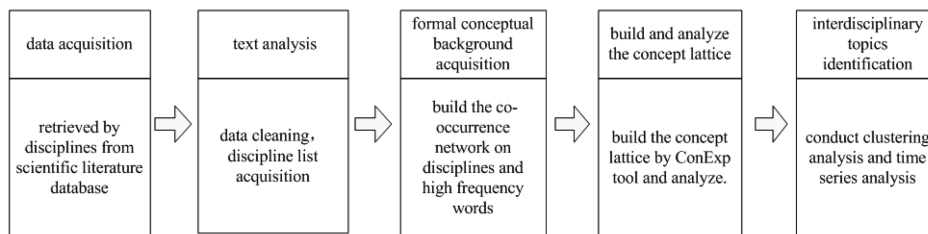Five main steps are conducted in the present study (Figure 1).

Figure 1.   Procedures for interdisciplinary knowledge discovery based on the CLT.

Step 1: Data acquisition. The raw analytical data is retrieved from Web of Science database and then the titles and abstracts of the articles are extracted and preprocessed.

Step 2: Text analysis. Here the topic clusters acquisition method proposed by Porter is adopted (Porter & Zhang, 2012). After preprocessing the data set based on the text content, meaningful phrases that passed a certain threshold could be selected as a candidate topics set. Subsequently, data preprocessing is implemented using the Derwent Data Analyzer (DDA). For massive amounts of scientific literature, mining and cleaning of terms are time-consuming and laborious processes. DDA is a professional text-mining software package developed by Search Technology Inc. and Georgia Tech (DDA, 2018).

When the raw data is input into the DDA, fields with relative higher coverage for analysis are selected in order to ensure that the data set is comprehensive. Next, cleaning tools and a similarity merge command are utilized based on the phrases term forms to preliminarily merge the phrases terms. In this step, numbers, stop words, and many irrelevant words (pan-words) will be removed. Data filtering and pre-processing is a key step in interdisciplinary analysis, and data cleaning can greatly reduce the size of vocabularies. However, it is necessary to avoid excessive cleaning because the characteristic words in an interdisciplinary field are often associated with multiple disciplines, and thus, they are associated with multiple keywords. Therefore, if the same approach is employed to clean and merge these words, some meaningful words may be omitted. Through this process, realistic IDR topics can be obtained, which formed the basis of the subsequent interdisciplinary trend prediction process.

Step 3: Formal conceptual background acquisition. High-frequency words in the research field were obtained, the co-occurrence network is built based on the disciplines and high frequency topic terms, and we treated this co-occurrence network as a formal conceptual background. The co-occurrence network can be obtained using DDA. According to the disciplines-topic terms, binary co-occurrence networks, the data structure obtained is shown in Table 1.

Table 1.   Formal background data structure based on the formal context.

	$Term_1$	$Term_2$	…	$Term_n$
DIS_1	A_{11}	A_{12}		A_{1n}
DIS_2	A_{21}	…		
…			…	…
DIS_m	A_{m1}			A_{mn}

DIS_m: disciplines, $Term_n$: high frequency topic terms.

Step 4: The building of concept lattice, where the ConExp 1.3 program (Serhiy, 2000) is used to build and analyze the concept lattice.

Step 5: Clustering analysis and time-series analysis for IDR topic identification. Clustering result can also be obtained by using the ConExp 1.3 program, using

which we could recognize interesting concepts. The time-series analysis can clarify the origin and development process in order to help observe the evolution trajectory of interdisciplinary knowledge and predict the future IDR topics.

5.2 Data acquisition and text analysis

In this study, the discipline classification refers to Web of Science Subject Categories, which is widely used discipline classification built by citation analysis and field expert instructions. Information Science & Library Science (LIS) is a highly interdisciplinary subject, which involves multiple disciplines in science, engineering, agriculture, medicine, management, economics, and other fields. We think three important factors promote the development of LIS, i.e. a solid foundation for theory, innovative techniques, and the expansion of application fields (Xu et al., 2015). In Web of Science, Clarivate Analytics defines LIS as an emerging interdiscipline: LIS covers a wide variety of research topics, including bibliographic study, cataloguing, categorization, database construction and maintenance, electronic libraries, information ethics, information processing and management, interlending, preservation, scientometrics, serial librarianship, and special libraries (Clarivate Analytics, 2017).

In the present study, we employed research publications from the field of LIS as the measurement data, via the following search strategy: WC = Information Science & Library Science, where we retrieved the articles from 2007 to 2016 on 2nd January, 2017, and 41,980 records were accessed. The overall data set was too large for a text analysis, so we selected data set of five odd-numbered years to analyze the IDR topics in the near ten years, which include 2007, 2009, 2011, 2013, and 2015. Term frequency is the number of times that a particular keyword appears in a document. In order to improve the visualization effect, top 60 high frequency terms were selected as basic data each year, and the concept background with the related disciplines were finlly formed.

5.3 Determination of the formal conceptual background and concept lattice

Table 2 shows the partial formal conceptual background in 2007, number "1" means the column term (topic terms) appeared in the corresponding row (discipline), while number "0" means there is no co-occurrence relationship between them. Figure 2 shows the concept lattice based on the knowledge structure for LIS in 2007, where a node represents a concept. A node with a blue semicircle shows that the concept has a connotation (topic terms), whereas a node with a black semicircle denotes that the concept has an extension (discipline). Disciplines are tagged with white labels and topic words are tagged with gray labels.

Research Paper

Table 2.   Formal conceptual background in 2007 (partial data).

	User's information needs	Information technology	Information retrieval	Citation data and analysis methods	Qualitative analysis
Computer Science-Information Systems	1	1	1	1	1
Computer Science-Interdisciplinary Applications	1	1	1	1	1
Management	1	1	1	1	1
Communication	1	1	0	1	1
Multidisciplinary Sciences	0	0	0	0	0
Medical Informatics	1	1	1	1	1
Geography	1	1	1	0	1
Geography-Physical	1	1	1	0	1
Social Sciences-Interdisciplinary	0	1	0	0	1
Telecommunications	1	0	0	1	1

The upper layer concept connotation (topic term) is inherited by the lower layer, so we can treat the connotation of the upper concept node as the main interdisciplinary research field. If the levels of the concept nodes are deeper, their connotations are richer and each research topic is kept on subdivided. Each related topic term can be described by the upper layer concept connotation connected to it, and each discipline related to topics can be determined by the concept extension. The topic connotation is broader if the concept layer of the topic is located higher, while the lower level of disciplines in a concept and the more number of IDR topics with LIS. If the discipline has more IDR topics, the discipline is located at a lower level. The discipline at the lowest level contains all of the topics that are connected to it.

According to the knowledge representation characteristics of concept lattice, different topic terms are clustered into groups and rendered as branches of the concept lattice, in other words, terms belong to same group are shown in a single branch. Therefore, the IDR topics are identified based on whether the branches directly connect with the top-level node in whole concept lattice, and each branch represents an IDR topic. Moreover, the relationship among different terms in the same topic is not merely relevance. It is an extension and inheriting of the semantic meaning, which is revealed by multiple links between the upper and lower layers. What's more, semantic distance of different IDR topics can be easily identified through counting the number of nodes covered by different topics. The greater the number of mutual nodes, the more close of their semantic similarity. If an IDR topic does not have mutual nodes with others, it can be speculated that its content and connotation is distinctive. In addition, it puts forward several new approaches to estimate the importance of IDR topics, including analyzing the location of its top-level node and measuring the number of nodes it covers. IDR topics with higher-level locations or more nodes are more broader connotation than the others.

Especially, the node on the highest level has the broadest knowledge connotation
and extension, so each topic is named by its respective top-level term. Finally, it
can specifically identify the most particular and distinct terms of a discipline,
namely, the terms appear in the same node with a specific discipline are the most
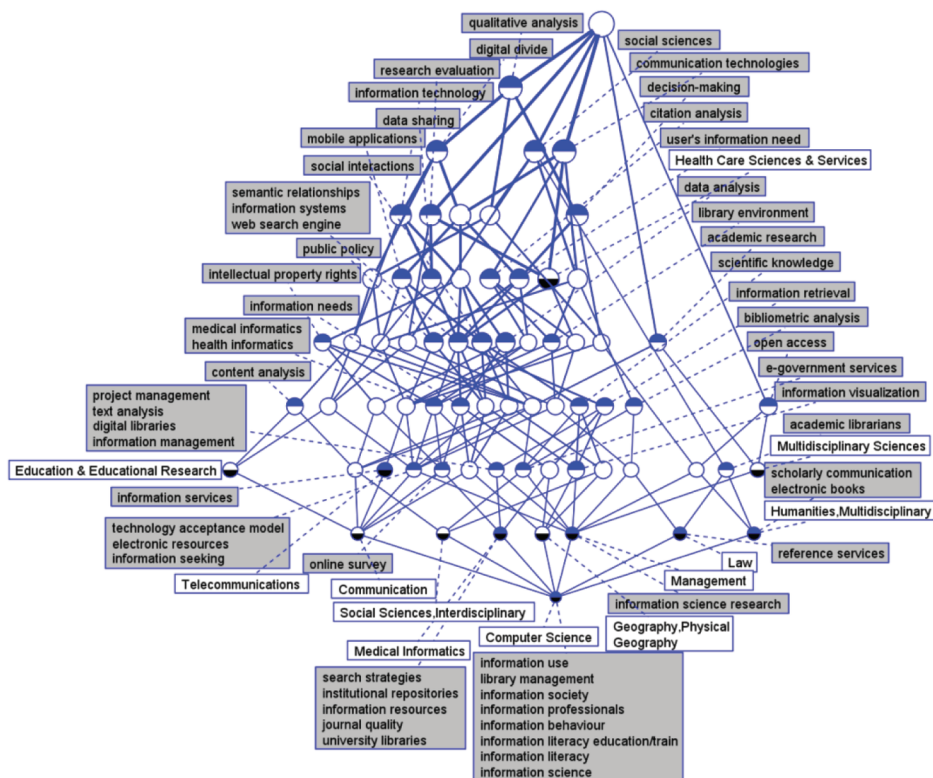distinctive ones of this discipline.

Figure 2.   Concept lattice in 2007.

From Figure 2, we can see all the disciplines and intersecting topics with LIS.
There are 14 disciplines that have IDR topics with LIS according the FCA rule.

To carry out microanalysis on IDR topics, we focus on the IDR topics and
evolution trends over time of specific disciplines, which include Medical Informatics
and Geography-Physical. According the formal conceptual background of each year,
it can be seen that there are much more IDR topics between LIS and Medical
Informatics than Geography-Physical. We try to respectively find the IDR topics
between LIS & Medical Informatics and LIS & Geography-Physical and take
a comparative analysis. On the basis of its theoretical interpretation, Medical
Informatics covers resources on health care information in clinical studies and

medical research. This category includes resources on the evaluation, assessment, and use of health care technology, its consequences for patients, and its impact on the society (Clarivate Analytics, 2017). Geography-Physical covers resources dealing with the differentiation of areas of the Earth's surface as shown in the character, arrangement, and interrelations over the world of such elements as climate, elevation, soil, vegetation, population, land use, industries, or states, as well as the unit areas formed by the complex of these individual elements (Clarivate Analytics, 2017).

6   Empirical Analysis

6.1   IDR topics identification

(1) IDR topics identification in 2007

Figure 3a is an IDR topics concept lattice between LIS and Medical Informatics in 2007. It covers three topics, including qualitative analysis, research evaluation, and communication technologies. Qualitative analysis is the most important research topic in the interdisciplinary field since it is located on the first level of the concept lattice. It has direct interaction with the lower level topic terms, which include digital divide and decision making. The three nodes and information technology, which are situated on a level further below are all unique nodes of qualitative analysis. Communication technologies are on the second level and only citation analysis directly interacts with it on the lower level. Both are unique nodes of the corresponding topics. Research evaluation is on the third level of the concept lattice and it is the only unique node within the topic. The three topics mentioned above have close semantic relation and share multiple topic terms like semantic relationships, information systems, web search engine, medical informatics, health informatics, information needs, project management, text analysis, digital libraries, information management, and online surveys. By hierarchical network structure of concept lattice, three topics relate to each other through their semantic extension and mutual cross. During the process of semantic extension, finally, they connect with the most distinct terms of Medical Informatics—search strategies.

Figure 3b shows the concept lattice of IDR topics between LIS and Geography-Physical in 2007. Three topics, qualitative analysis, research evaluation and communication technologies, are contained. These topics are consistent with the IDR topics between LIS and Medical Informatics, meanwhile, the importance of the topics and their unique topic terms are identical. Nonetheless, the terms for sharing of the topic changed. All topic terms except semantic relationships, information systems, web search engine, and information needs no longer belong to sharing
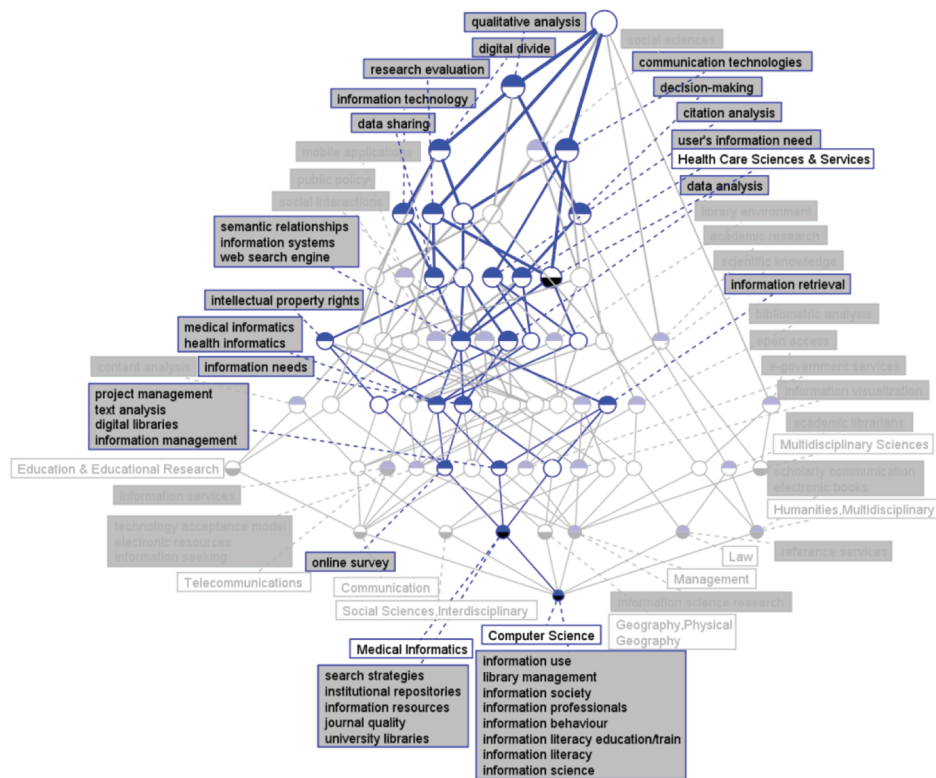
Figure 3a.   IDR topics in concept lattice between LIS and Medical Informatics in 2007.

topic terms. What's more, there are some new sharing topic terms, including information retrieval and information visualization. Thus, the IDR topics connotation in 2007 of LIS interacting with Geography-Physical is not completely equivalent with the one with Medical informatics. Furthermore, the semantic distance between qualitative analysis and communication technologies is closer. In addition, the sharing nodes of two topics account for more than half of the whole concept nodes, such as the user information need, public policy, social interactions, and data analysis.

(2) IDR topics identification in 2009

Figure 4a presents IDR topics concept lattice between LIS and Medical Informatics in 2009 and it contains four IDR topics, including qualitative analysis, communication technologies, health informatics, and content analysis. Communication technologies are on the first level of the concept lattice and the nodes it covers are far more than other topics. Besides, communication technologies has the most rich connotation and maximum unique nodes, such as digital divide, data analysis, information
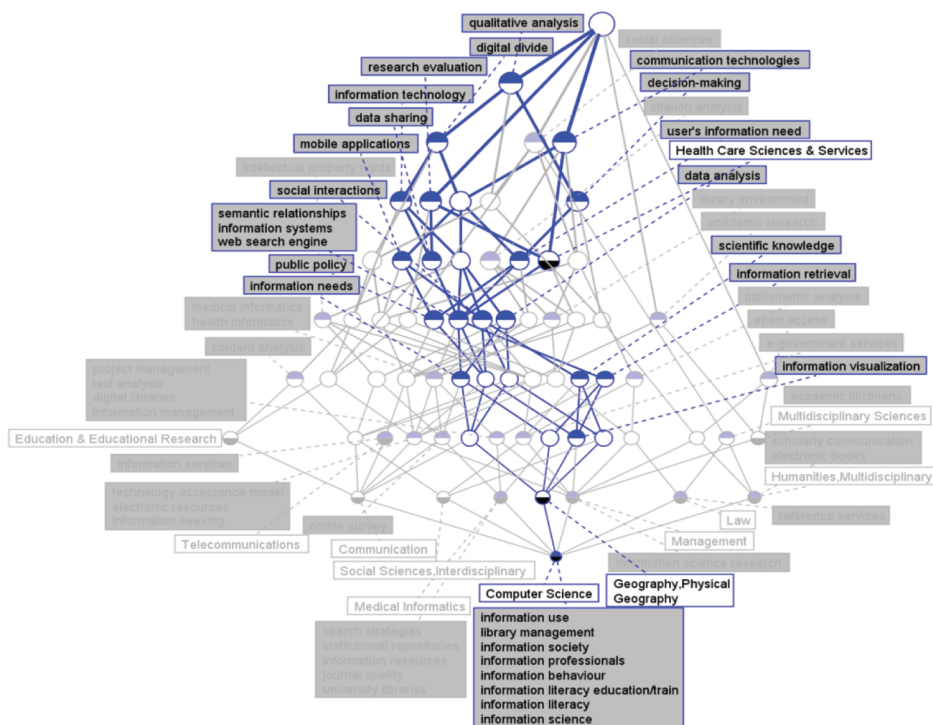
Research Paper

Figure 3b.   IDR topics in concept lattice between LIS and Geography, Physical for 2007.

technology, semantic relationships analysis, data collection, project management, user satisfaction, and information management. By contrast, the richness of semantic connotation of the other three topics is relatively low. Moreover, compared with 2007, the number of sharing nodes of all the topics, consisting of interdisciplinary research, information needs, metadata, information retrieval, and search engines is decreased, indicating that the semantic relevance of all topics declined.

Figure 4b is the IDR topic concept lattice of LIS with Geography-Physical in 2009. It contains three IDR topics, which include qualitative analysis, communication technologies, and web2.0. Comparing with the IDR topics between LIS and Medical Informatics in 2007, health informatics, and content analysis disappear while web2.0 emerges. However, it is constant that communication technologies is still on the first level of the concept lattice, and this topic covers more than half the nodes. Besides, it has the maximum unique nodes and the richest semantic connotation. In 2009, the semantic relevance among three topics significantly reduces since there are no topic terms covered by all topics. Only decision making, information sharing, information science, and information systems cover two of them.

Figure 4a.   IDR topics in concept lattice between LIS and Medical Informatics for 2009.

Figure 4b.   IDR topics in concept lattice between LIS and Geography-Physical for 2009.

Research Paper

(3) IDR topic identification in 2011

Figure 5a presents IDR topics concept lattice between LIS and Medical Informatics in 2011. It involves three IDR topics, which are information technology, digital divide, and open access. Digital divide is situated on the first level and it covers all topic terms except information technology and open access. Therefore, this topic plays an important role in the interdisciplinary field. The nodes on the lower level that directly interact with the digital divide include information system, health information, and user acceptance. In addition, the nodes on the lower level, which directly interact with these nodes, involve data collection and perceived usefulness. These and the nodes on the lower level such as technology acceptance model, user satisfaction, information needs, information seeking and information seeking behavior are all unique nodes of digital divide. In addition, the semantic relevance among three topics is relatively weak. Only citation counts and information services are covered by digital divide and open access simultaneously, and all the other nodes are unique nodes of corresponding topics. What's more, information retrieval is the most distinct term of MedicalInformatics in 2011.
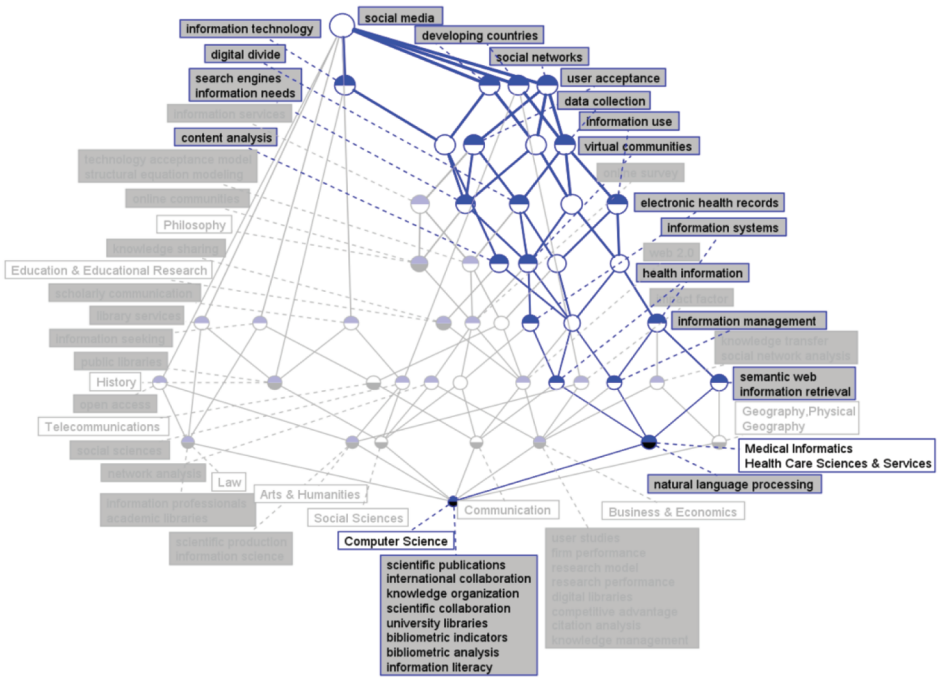
Figure 5a.   IDR topics in concept lattice between LIS and Medical Informatics for 2011.

Figure 5b is the IDR topic concept lattice of LIS with Geography-Physical in 2011. It includes three IDR topics, which are social networks, information technology, and social sciences. They are very different from the IDR topics between LIS and Medical Informatics. Among these, social networks and information technology are both located on the first level of concept lattice and they are more important than social sciences. In the whole concept lattice, nodes are scarce and covers only four topic terms, and at the same time, only social networks has a node in the lower level, indicating that the semantic connotation of three topics is not rich enough. What's more, no topic term is covered by all topics, which suggests that semantic relevance of these topics is not intimate.
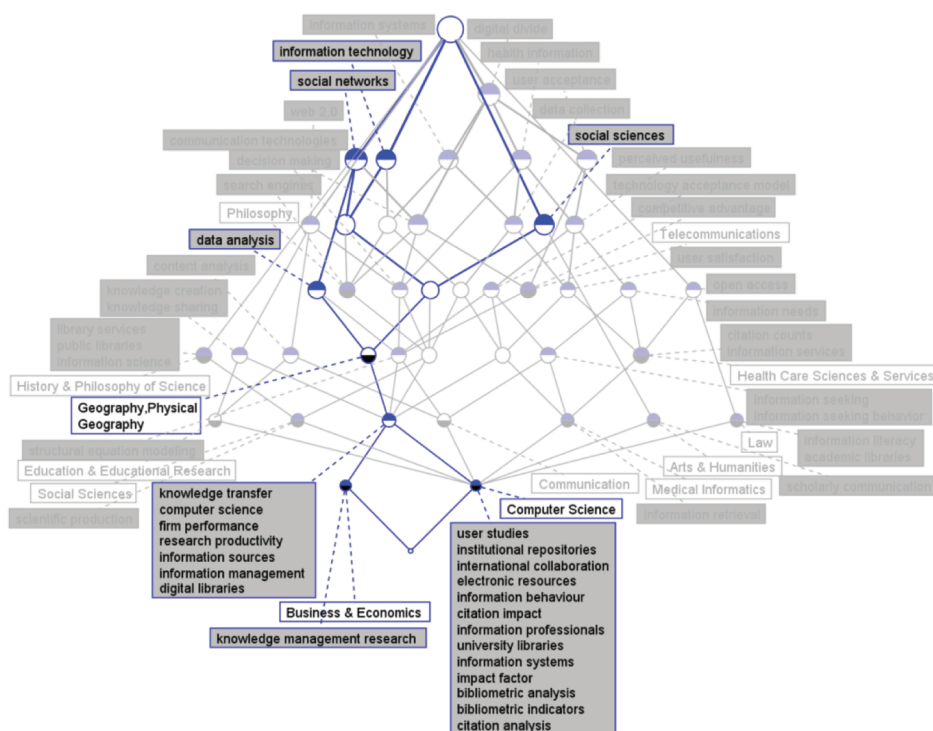
Figure 5b.   IDR topics in concept lattice between LIS and Geography-Physical for 2011.

(4) IDR topic identification in 2013

The IDR topic concept lattice between LIS and Medical Informatics in 2013 is shown in Figure 6a. There exist four IDR topics, including information technology, social media, developing countries, and social networks. There is no difference in the importance of four topics from the perspective of the top node's location.

Research Paper

However, social networks have the most important status from the perspective of quantity of nodes. In the lower level, user acceptance and data collection directly interact with social networks, and user acceptance is the unique node of social networks. Overall, there are quite few unique nodes, and all topics only have one unique node except social networks, which has two unique nodes. Most nodes are covered by multiple topics and virtual communities, electronic health records, health information, information management, and natural language processing are shared by four topics, implying that the semantic relevance of the IDR topics in 2013 is very close. In addition, natural language processing is the most distinct term of Medical Informatics in 2013.

Figure 6a.   IDR topics in concept lattice between LIS and Medical Informatics for 2013.

Figure 6b is the IDR topics concept lattice of LIS interacting with Geography-Physical in 2013. It has three IDR topics, which involve social media, developing countries, and social networks. These topics are all included in the IDR topics between LIS and Medical Informatics. The top nodes of the corresponding topics are all on the first level of the concept lattice. However, the number of topic terms included in each topic is significantly different. There into, social networks almost covers all topic terms, demonstrating that its semantic connotation is relatively rich.

However, in general, the amount of topic nodes is fewer in the concept lattice and the trend of interdisciplinarity is not obvious. Among three topics, the semantic relevance between developing countries and social networks is closer, as all topic terms are covered by them except the top-level nodes. These topic terms are data collection, information use, information systems, knowledge transfer, social network analysis, semantic web, and information retrieval.
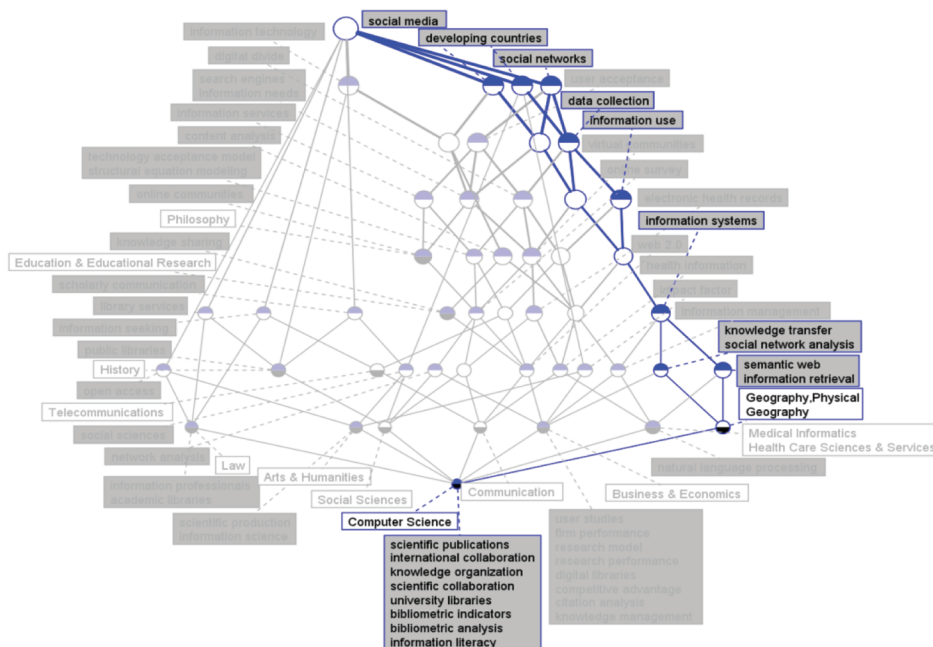

Figure 6b.   IDR topics in concept lattice between LIS and Geography-Physical for 2013.

(5) IDR topics identification in 2015

As shown in Figure 7a, the IDR topics concept lattice between LIS and Medical Informatics in 2015 involves seven IDR topics, which include social media, decision making, knowledge sharing, data analysis, information system, mobile applications, and qualitative analysis. Social media, data analysis, mobile applications, and qualitative analysis are all on the first level of the concept lattice, suggesting highest importance. Compared to other years, the concept lattice in 2015 has the most topic terms and its association relationship is quite intricate and complex, signifying that the interdisciplinary content of two disciplines is quite rich. In the view of the semantic relationship among the topics, social media has unique nodes, indicating that it has a unique semantic meaning. Data analysis and qualitative analysis have most common nodes, implying that the semantic association of them is most closely

Research Paper

related. Besides, the nodes that are shared by more than four topics include semi-structured interviews, decision-making support, technology acceptance model, search engine, perceived usefulness, and user acceptance, which can be regarded as important hinge nodes in interdisciplinary filed in 2015. In addition, natural language, electronic health record, and information retrieval are the most distinct terms of Medical Informatics in 2015.
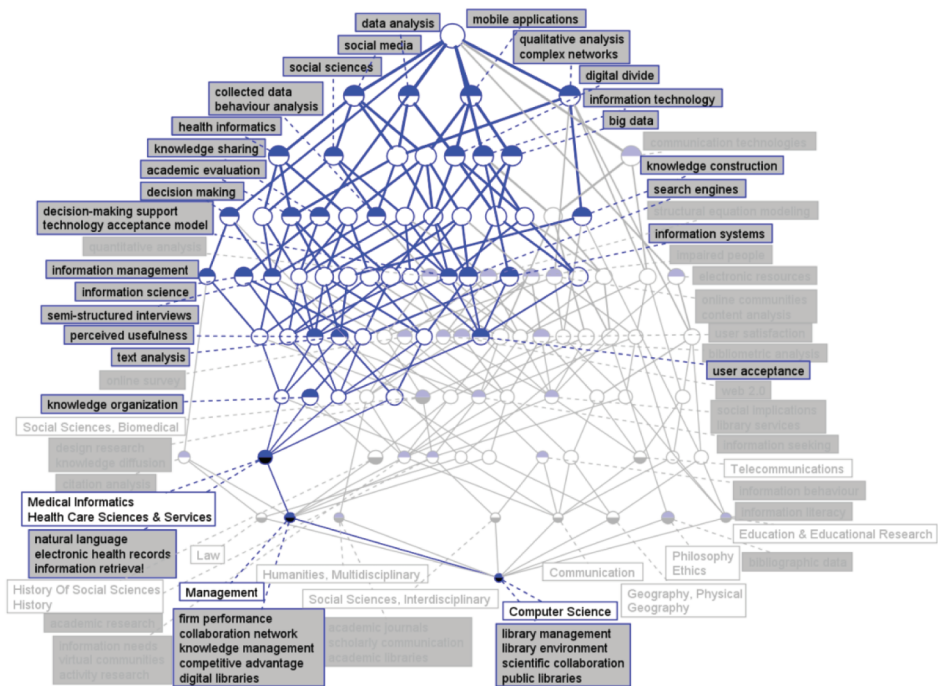

Figure 7a.   IDR topics in concept lattice between LIS and Medical Informatics for 2015.

As can be seen from Figure 7b, the IDR topics concept lattice of LIS interacting with Geography-Physical in 2015 involves five IDR topics, which contain social media, data analysis, mobile applications, electronic resources, and qualitative analysis. Decision making, knowledge sharing and information system are not present as in the IDR topics between LIS and Medical Informatics in 2015. All topics except electronic resource are equally important, as the top nodes of corresponding topics are all on the first level of the concept lattice. Electronic resources has no sharing topic terms with other topics, demonstrating that the semantic relationship between it and other topics is quite far. The semantic relationship between data analysis and mobile applications is the closest since all the non-exclusive topic terms are shared by them except text analysis. These topic terms include search engines, impaired people, web 2.0, and user satisfaction.
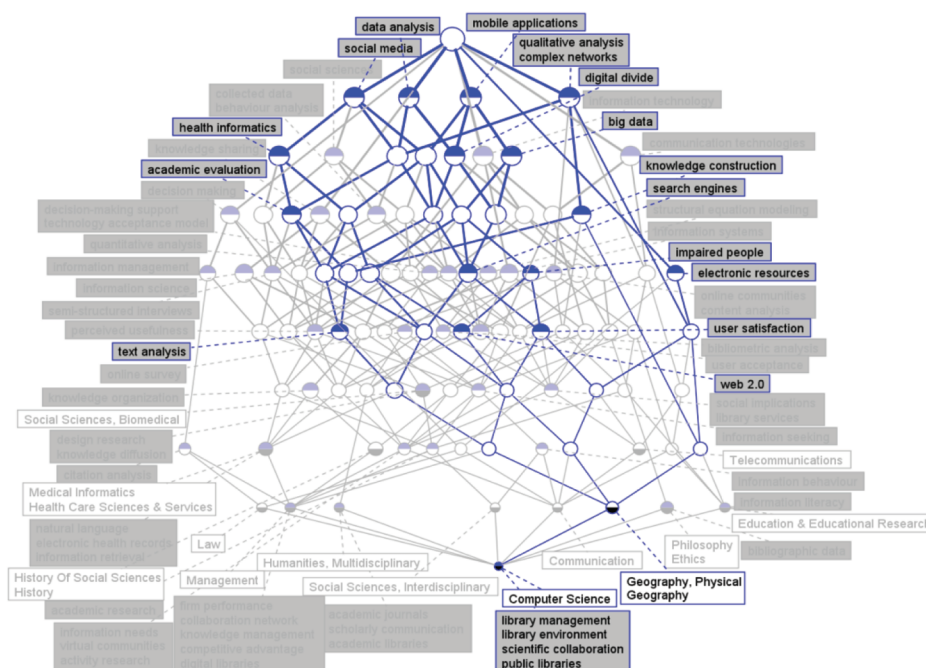
Figure 7b.   IDR topics in concept lattice between LIS and Geography-Physical in 2015.

6.2   Time series analysis and forecasting of IDR topics

a. Evolution of IDR topics

In this study, the IDR topic identification is within the disciplines of LIS, therefore, the result is more likely to reflect when it is applied in other disciplines. The following are the results of knowledge application from LIS to Medical Informatics and Geography-Physical.

In Medical Informatics, it is found that before 2011, the IDR topics of LIS and Medical Informatics are mainly concerned with several information analysis methods such as qualitative analysis, research evaluation and communication technologies, health informatics, and content analysis. After 2011, there were developments in both research objects and methods of LIS. On the basis of the basic IDR topics such as information technology, information system, data analysis, and qualitative analysis, IDR topics of LIS and Medical Informatics tend to present the characteristics of the problem-oriented, that is, taking methods of LIS to study or solve specific problems, including management and service provision in such areas as social media, social networks, and mobile applications. At the same time, the IDR topics are more concerned about the decision-making methods and applications

based on scientific data analysis, knowledge sharing, health care information, and medical problem in the developing countries.

For Geography-Physical, it is found that before 2011, the IDR topics of LIS and Geography-Physical were mainly concerned with the application of several information analysis methods in Geography-Physical, such as qualitative analysis, research evaluation, communication technologies, and web2.0. After 2011, the focus was more on management and use the massive electronic resources in Geography-Physical, mainly involving the information analysis method and tools of LIS, such as social networks, social sciences, information technology, and qualitative analysis. Meanwhile the IDR topic concerned more about the geographical resource in developing countries and application of social media in resources displaying.

It can be seen from the above analysis, although there are several similar IDR topics of LIS & Medical Informatics and LIS & Geography-Physical, the richness and content of these topics have differences. Obviously, compared to Geography-Physical, Medical Informatics has richer connotations of IDR topics with LIS.

b. Evolution of interdisciplinary research

After comparing the cross-disciplines related to Medical Informatics and Geography-Physical within the field of LIS from 2007 to 2016, it can be found that Computer Science (especially the areas of information systems and interdisciplinary applications) had relatively closer cross-correlations with both Medical Informatics and Geography-Physical. Computer Science is situated at a lower level than Medical Informatics and Geography-Physical, which means that Computer Science is an applied science with more common content attributes within LIS than Medical Informatics and Geography-Physical. In addition, Computer Science was the closest related discipline to LIS, which has become the main technique support for Medical Informatics and Geography-Physical, and it is important in the knowledge spillover process. Therefore, Computer Science plays an important role in the interdisciplinary applications of Medical Informatics and Geography-Physical, as well as in the construction of medical and geography information systems.

In the Medical Informatics discipline, Health Care Sciences & Services should receive more attention. "Health Care Sciences & Services cover health services and learning resources, hospital management, health care management, health care financing, health policy and planning, health economics, health education, medical history, and other types of palliative care research" (Clarivate Analytics, 2017). Its level in the concept lattice gradually moved down to be close to Medical Informatics and they were in the same concept node in 2013, which means Medical Informatics and Health Care Sciences & Services had the same IDR topics with LIS. It also suggests that more common research topics appear in the interdisciplinary area of Health Care Sciences & Services and LIS.

For the Geography-Physical discipline, the time series analysis showed that Geography should receive more attention, for it is always located in the same concept node. "Geography covers resources concerned with socio-cultural aspects of the Earth's surface emphasizing the human, economic, political, urban, and environmental issues of the discipline. The history of geography and the study of cartography are also covered in this category" (Clarivate Analytics, 2017). Although Geography focuses more on the socio-cultural aspects of the geographical study, it has the similar characteristics of IDR topics with LIS.

c. Prediction of interdisciplinary and IDR topics

In the Medical Information discipline, medical knowledge, decision support, information retrieval, hospital information systems, social outcomes evaluation, education, training, and other medical informatics research content still require the technical support of computer technology and information analysis methods. Thus, information analysis methods, computer technology, communication technology, social networks, social media are still important IDR topics for LIS and Medical Information research in the future.

For Geography-Physical, with the arrival of the big data era, various types of geographical data are increasing fast. Therefore, in the future, various methods of LIS for big data analysis will be important IDR topics for LIS and Geography-Physical.

Beyond that, Zhang et al. proposed the concept of "Subject Informatics" and noted that in the data-intensive scientific research paradigm, scientific research has gradually become a data-driven knowledge discovery activity, and thus, the era of data-driven science is emerging. Subject-specific areas of informatics based on data analysis have developed rapidly and they are applied widely. Meanwhile, the general knowledge system that supports information analysis and the application of subject-specific informatics has also been improved, thereby providing solid foundations for general subject informatics (Zhang & Fan, 2015). Thus, in this context, as a discipline for providing decision support based on data analysis, LIS can provide more effective analytical methods for use in Medical Informatics and Geography-Physical.

However, Medical Informatics is a typical form of Subject Informatics that will develop into a discipline that supports a new paradigm in LIS research. However, Geography-Physical may still have some way to go before being a Subject Informatics, for there is not a specific Subject Informatics for Geography but Geo-spatial information science has existed for years, and provides a certain foundation to formation Subject Informatics. In the future, the interdisciplinary relationships among Medical Informatics, LIS, and Health Care Sciences & Services will be

further strengthened as their content gets deeply integrated. Meanwhile, an increasing number of intelligence analysis methods and tools will be absorbed in Geography-Physical studies, which gradually will help it grow and expand into a powerful Subject Informatics.

6.3 Comparative analysis and discussion

To evaluate the advantages and effectiveness of topic recognition based on the FCA method, we further performed a comparative analysis with two other IDR topic recognition methods. Both methods employ LIS as the empirical field, and hence, through comparison, we can examine the advantages and disadvantages of the IDR topic recognition method based on CLT. The first method involves an index series named terms interdisciplinarity (TI index), which attempts to recognize topics by calculating TI values together with Bet values and term frequency values and analyzes the evolution of interdisciplinary sciences based on social network analysis and time-series analysis. A study has proved that the TI value can identify IDR topic terms well (Xu et al.,2016). The second method is an integrated method for IDR topic recognition and prediction, which integrates various methods, including co-occurrence networks analysis, high-TI terms analysis, and burst detection, and offers an overall perspective into interdisciplinary topic identification (Dong et al., 2018).

Through the comparative analysis, we concluded that IDR topic recognition based on the CLT has its own advantages compared with the other methods that use topic terms.

First, the IDR topic recognition based on the CLT can easily discover a specific IDR topic, which is generally located at the lower part of the concept lattice. Therefore, for the interpretation and prediction of IDR topics, further reference of the upper semantic nodes and detailed IDR topic recognition and prediction are necessary. For example, in 2009, for communication technology, topic terms located in the lower part can be interpreted as follows. With the continuous development of communication technology, the level of information reception of different social groups, countries, and regions gradually produces a gap, which is called a digital divide. Eliminating the digital divide is imperative, and it can be achieved by increasing the opportunity and methods of data collection for the weaker side, understanding user satisfaction, and strengthening user management and information management. Moreover, adding an information communication path through the online community can help eliminate the digital divide. In this process, the use of interdisciplinary approaches is required, and simultaneously, the user information should be subjected to orientation, text analysis, and metadata processing as basic technologies to further enhance the recall and precision of information retrieval and

search engine optimization. This is an effective way by which information resources can be fully activated and utilized.

Second, the topic directly connected to the top node tends to be the traditional and important topic, whereas the convergent nodes are more targeted and specific IDR topics and furthermore are the future direction of IDR. For example, in 2009, communication technology was a stable and significant topic term as it was located at the first level, and information sharing tended to represent the future trend of IDR topics as several topics converged to it.

Furthermore, the bipartite network only shows the co-occurrence relation between the discipline and the topic terms. The different hierarchies of the concept lattice contain the change process of the semantic relationship between the disciplines and the topic terms. For example, in 2013, the closeness of the semantic relationship between developing countries and social networks could not be measured in the dichotomous networks. However, their close connection was demonstrated through numerical analysis of mutual nodes in the concept lattice.

Although the IDR topic recognition based on the CLT has advantages, it has its own limitations compared with the other methods. For example, IDR topic recognition based on the CLT is not sensitive to the terms interdisciplinarity in contrast to the TI index. Furthermore, it cannot recognize IDR topics from different perspectives in contrast to the integrated method.

7 Conclusion

In this study, we introduced the concept and application of FCA and CLT, proposed an IDR topics discovery method based on the CLT, and performed an empirical analysis on the IDR topics of LIS & Medical Informatics and LIS & Geography-Physical with a comparative analysis. Further, we proposed two advantages in the identification and prediction of IDR topics based on the CLT. The empirical findings proved that the concept lattice approach is suitable for IDR topic identification and predictions.

Multiple advantages of CLT in knowledge representation make it an important IDR topic identification method. However, there are still limitations when using CLT to discover the IDR topics. In a concept lattice, the data can only reflect whether there is a relationship between the discipline and the topic terms, and hence, they will fail to represent the strength of the relationship between them. Moreover, the CLT cannot clearly represent more concepts. Therefore, the number of topic terms analyzed in this study was limited. The crucial issue towards the discovery of knowledge using CLT is knowledge reduction of knowledge representation while maintaining structure consistency. Considering the rapid improvement of both

Research Paper

complex network analysis and topic semantics analysis, in the future, it is possible to combine concept lattice analysis and complex network analysis to obtain a multi-level relationship, which can be used to identify IDR topics.

Acknowledgements

Author contributions

Haiyun Xu (xuhy@clas.ac.cn, corresponding author) proposed the research idea, designed the research, drafted and revised the manuscript. Chao Wang (wangchao1@sdas.org) performed the data processing, and revised the manuscript. Kun Dong (dongkun@sdut.edu.cn) performed the research and revised the manuscript. Zenghui Yue (yzh66123@126.com) performed the research and revised the manuscript.

References

Belohlavek, R., & Vychodil, V. (2009). Formal concept analysis with background knowledge: Attribute priorities. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 39(4), 399–409.

Carpineto, C., & Romano, G. (2004). Concept data analysis: Theory and applications. John Wiley & Sons.

Chang, Y.W., & Huang, M.H. (2012). A study of the evolution of interdisciplinarity in library and information science: Using three bibliometric methods. Journal of the American Society for Information Science and Technology, 63(1), 22–33.

Cimiano, P., Hotho, A., & Staab, S. (2005). Learning concept hierarchies from text corpora using formal concept analysis. Journal of Artificial Intelligence Research, 24, 305–339.

Dong, K., Xu, H., Luo, R., Wei, L., & Fang, S. (2018). An integrated method for interdisciplinary topic identification and prediction: A case study on information science and library science. Scientometrics, 115(2), 849–868.

Ganter, B., & Wille, R. (2012). Formal concept analysis: Mathematical foundations. Springer Science & Business Media.

Ganter, B., & Wille, R. (1997). Applied lattice theory: Formal concept analysis. In General Lattice Theory, G. Grätzer editor, Birkhäuser.

Gao, J.F. (2015). Coupling network literature knowledge discovery based on concept lattice. Library science research, 56(17), 122–125.

Hammarfelt, B. (2011). Interdisciplinarity and the intellectual base of literature studies: Citation
    analysis of highly cited monographs. Scientometrics, 86(3), 705–725.

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between
    two documents. Journal of the American Society for information Science, 24(4), 265–269.

Jia, C.Y., & Ni, X.J. (2003). Association rule mining: A survey. Computer Science, 30(4), 145–148.

Klein, J.T. (2000). A conceptual vocabulary of interdisciplinary science. Practising interdisciplinar-
    ity, 3–24.

Kumar, C. (2011). Knowledge discovery in data using formal concept analysis and random
    projections. International Journal of Applied Mathematics and Computer Science, 21(4),
    745–756.

Lahcen, B., & Kwuida, L. (2010). Lattice miner: A tool for concept lattice construction and explo-
    ration. Suplementary Proceeding of International Conference on Formal concept analysis
    (ICFCA'10).

Leydesdorff, L., & Rafols, I. (2011). Indicators of the interdisciplinarity of journals: Diversity,
    centrality, and citations. Journal of Informetrics, 5(1), 87–100.

Leydesdorff, L., Rafols, I., & Chen, C. (2013). Interactive overlays of journals and the measure-
    ment of interdisciplinarity on the basis of aggregated journal–journal citations. Journal of the
    American Society for Information Science and Technology, 64(12), 2573–2586.

Li, C.L., Liu, F.F., & Guo, F.J. (2013). Analysis on interdisciplinary research topics with cfinder
    of overlapping communities visualization software—taking the information science and
    computer science for example. Library and Information Service, 57(7), 75–80.

Liu, P., & Wang, Z. (2012). A new method for detecting or ganizational knowledge structure:
    Author keyword coupling analysis based on FCA. Library and Information Service, 56(22),
    121–128.

Liu, P., & Wu, Q. (2014). Detecting disciplinary knowledge structure based on formal concept
    analysis: An empirical investigation on library and information science, 58(18), 50–65.

Min, C., & Sun, J.J. (2014). Clustering analysis on discipline-crossing research hotspots: An
    example of library and information science and journalism and communication studies.
    Library and Information Service, 58(1), 109–116.

Porter, A.L., Roessner, J.D., & Heberger, A.E. (2008). How interdisciplinary is a given body of
    research. Research Evaluation, 17(4), 273–282.

Porter A, Zhang Y. Text clumping for technical intelligence. Theory & Applications for Advanced
    Text Mining, 2012.

Reuters, T. (2016). Science citation index expanded. http://ip-science.thomsonreuters.com/mjl/
    scope/scope_scie/.

Schummer, J. (2004). Multidisciplinarity, interdisciplinarity and patterns of research collaboration
    in nanoscience and nanotechnology. Scientometrics, 59(3), 425–465.

Shao, Z.Y., & Li, X.X. (2015). Detecting interdisciplinary knowledge structure based on concept
    lattice and bibliographic coupling. Library and Information Service, 59(8), 78–86.

Stumme, G. (2009). Formal concept analysis. In: Staab S., Studer R. (eds) Handbook on Ontolo-
    gies. Springer Berlin Heidelberg, 177–199.

Teng, G.Q. (2012). Research on knowledge organization based on concept lattice of digital library.
    Changchun: Jilin University.

Teng, G.Q., & Bi, Q. (2010). Comparative study on ConExp and lattice miner. New Technology of Library and Information Service, 26(10), 17–22.

Teng, G.Q, Bi, Q., & Bao, Y.L. (2011). An analysis on keywords of literature based on granularity concept analysis—A case study of ontology. New Technology of Library and Information Service. 27(9), 1–6.

Derwent Data Analyzer. (2018). Retrieved from https://www.thevantagepoint.com/tda-home.html

Venter, F.J., Oosthuizen, G.D., & Roos, J.D. (1997). Knowledge discovery in databases using lattices. Expert Systems With Applications, 13(4), 259–264.

Wille, R. (2002). Why can concept lattices support knowledge discovery in databases? Journal of Experimental & Theoretical Artificial Intelligence, 14(2–3), 81–92.

Wille, R. (2009). Restructuring lattice theory: An approach based on hierarchies of concepts. Formal Concept Analysis. Springer Berlin Heidelberg.

Xu, H.Y., Guo, T., Yue, Z.H., Ru, L.J., & Fang, S. (2016). Interdisciplinary topics of information science: A study based on the terms interdisciplinarity index series. Scientometrics, 106(2), 583–601.

Xu, H.Y., Liu, C.J., Lei, B.X., Li, H.L., & Fang, S. (2014). Measurement visualization and application of interdisciplinary research. Library and Information Service, 58(12), 95–101.

Xu, H.Y., Yin, C.X., Guo, T., Tan, X., & Fang, S. (2015). Interdisciplinary research review. Library and Information Service, 59(5), 119–127.

Serhiy, A. Yevtushenko (2000). System of data analysis "Concept Explorer". Proceedings of the 7th national conference on Artificial Intelligence KII-2000, p. 127–134.

Zhang, H.L., Wei, J.X., Du, Z.D., Liu, X., YAN, S., Feng, Z., Li, X.D., & Feng, X.F. (2011). Interdisciplinary research based on social complex network. Journal of Intelligence, 30(10), 25–29.

Zhang, Z.Q., & Fan, S.P. (2015). On the emergence and development of subject informatics. Journal of The China Society for Scientific and Technical Information, 34(10), 1011–1023.