How were science mapping tools applied? The Application of science mapping tools in LIS and non-LIS domains

Jiangen He
Department of Information Science, Drexel University
Philadelphia, United States
jiangen.he@drexel.edu

Wen Lou
Department of Information Management, East China Normal University, China
wlou@infor.ecnu.edu.cn

Kai Li
Department of Information Science, Drexel University
Philadelphia, United States
kl696@drexel.edu

ABSTRACT
As a scientific field, scientific mapping offers a set of standardized methods and tools that can be consistently adopted by researchers in different knowledge domains to answer their own research questions. This study examined scientific articles that applied science mapping tools to analyze scientific domains and the citations of these application articles. To understand the roles of these application articles in scholarly communication, we analyzed them and their citations by classifying them into library and information science (LIS) and other fields (non-LIS) in terms of both publication venues and analyzed domains. Through our preliminary study, we found that science mapping, a topic that is deeply situated in the LIS field, has gained increasing attention from various non-LIS scientific fields over the past few years. The application articles within and outside of the LIS fields played different roles in advancing the application of science mapping.

KEYWORDS
Science mapping tools; Scholarly communication; Scientific software; Scientific evolution.

ASIS&T THESAURUS
Citation analysis; Scientometrics; Citation impact.

INTRODUCTION
Software is the engine of the contemporary scientific system. It covers almost every aspect of the scientific practice that researchers are conducting on a daily basis in the increasingly more data-driven environment (Hey, Tansley, Tolle, & others, 2009). While data analysis is the focus of the majority of scientific software, many of software objects in scientific contexts are dedicated to represent scientific data and results in visual manners.

These visualization tools play important roles in the production of scientific knowledge, often more than they are credited. One of such examples is that they are a central part of the mechanism of science as “inscription device”, i.e., the function to create inscriptions that can be used for scientific arguments (Latour & Woolgar, 1979). Another reason for the strong rhetorical abilities of scientific graphs are also reflected in their capacity as immutable mobile (Latour, 1987), i.e., objects that can maintain their identities and meanings across time and space. This is also consistent with the idea that scientific visualization method is a boundary object (Star & Griesemer, 1989) across different knowledge domains. As a highly standardized set of method and tools, scientific visualizations can be easily used on different research problems by researchers from various communities.

One type of scientific visualization that is strongly connected to the field of information science is science mapping. Science mapping is a generic process of domain analysis and visualization (Chen, 2017). The process utilizes various methods and techniques including scientometrics methods (Callon, Courtial, & Laville, 1991; Kessler, 1963; Small, 1973; White & McCain, 1998), visualization techniques (Chen, 2006; Herman, Melancon, & Marshall, 2000; Shneiderman, 1996), and research metrics (Hirsch, 2005). The data used for the analysis process are scientific publications reflecting the scientific knowledge of the analyzed domain, which are usually collected from scholarly databases, such as the Web of Science, Google Scholar, and Scopus. Thanks to the availability of methods, techniques, and data, various science mapping tools (SMTs) have been developed to facilitate the process. These tools have become a necessary component in the science mapping study.

The method and tools of scientific mapping are an important site to understand how the material and methodological practice of science are represented in literary outputs of science, which in return, moderates how science is communicated. Several studies have examined the role of science mapping in scholarly communication by literature review (Chen, 2017), comparing science mapping tools (Cobo, López-Herrera, Herrera-Viedma, & Herrera, 2011), and analyzing the usage of science mapping tools (Pan, Yan, Cui, & Hua, 2018). However, the domain analyzed by using SMTs in scientific articles and the impact of science mapping studies that applied SMTs were rarely involved.

Understanding which domains have been analyzed in articles applied SMTs along with where these studies published can

82nd Annual Meeting of the Association for Information Science & Technology | Melbourne, Australia| 19 – 23 October, 2019

provide insights of the evolutionary patterns of science mapping. The citation analysis of these application articles can reveal the role of application articles for overcoming the domain boundary between Library and Information Science (LIS) and non-LIS domains. In this study, we examined the different kinds of SMT application articles in terms of their analyzed domains and journal domains and mapped out the citation impact of application articles. Specifically, the following research questions are pursued in this work-in-progress paper.

What is the evolutionary pattern of SMT application articles in LIS domains and non-LIS domains?

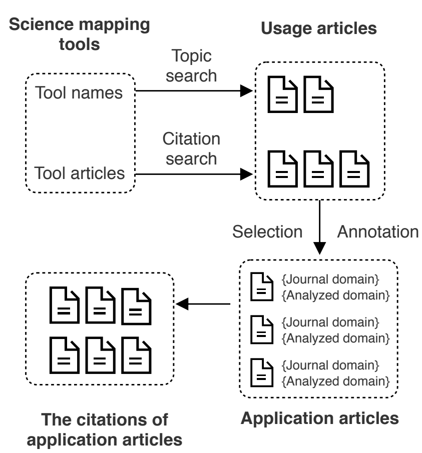
What are the scientific impacts of the application articles?

RESEARCH DESIGN

Science Mapping Tools and Articles

As described in Figure 1, we identified fourteen scientific mapping tools from literature (Cobo et al., 2011; Pan et al., 2018), after excluding ones which are not fit into our SMT definition, i.e., (1) data sources are scientific literature and (2) aim to understand and visualize the evolution of scientific science domains in whose original vision description. They are BiblioTool, CiteSpace, CitNetExplorer, CoPalRed, CRExplorer, HistCite, IDR Toolkit, ITGInsight, NEViewer, SciCurve, Sci2 Tool, SciMAT, VOSViewer, and VxInsight. Seventeen articles that are most highly cited individually and for establishing the SMTs were selected as tool articles.

Figure 1. The procedure of data collection.



928 distinct articles were collected that mentioned or cited all fourteen SMTs through both citation and topic search in the Web of Science. They are classified as usage articles.

This study only focuses on SMTs being actually used in scientific research but not all the papers citing/mentioning these tools. Identifying articles (*application articles*) that used these tools to 1) model and analyze structure and dynamic of scientific collaborations, or 2) measure and analyze scientific contributions and impacts (Chen, 2017) is a must. Application articles were identified manually. We downloaded and read the abstracts (and full texts if necessary) of 928 articles and two collaborators asset if a tool is used to report scientific research independently. Cohen's kappa coefficient was used to measure the inter-coder reliability (IRR) and the result is 0.929 (p=0.000). The cases of disagreement were assigned to the third collaborator. At last, 496 application articles were identified.

Annotating Application Articles

Researchers are more likely to use SMTs to conduct domain analysis. The applications in different domains, e.g., LIS and non-LIS domains provide different insights of the evolving role of science mapping as tools. Besides, the domains of journals where application articles published is another necessary dimension for our analysis.

Journal domains were identified by subject category field in data of Web of Science Core Collection. Journals that in the subject category "information science & library science" were annotated as "LIS journal", otherwise "non-LIS" journal. However, research domains in application articles analyzed by SMT cannot be determined by the subject category of journals. For example, a paper published in Scientometrics, a LIS journal may analyze nanotechnology or other non-LIS domains. Two collaborators manually identified domains studied in articles and then coded the domains as LIS or non-LIS domains. Two collaborators achieved an IRR of 0.951 (p=0.000). The cases of disagreement were resolved by the third collaborator.

Four categories were annotated as below, which are combinations of LIS domains (LD), non-LIS domains (ND), LIS journals (LJ), and non-LIS journals (LJ).

- Articles studied a **L**IS **d**omain and published in a **L**IS **j**ournal (LD-LJ);
- Articles studied a **L**IS **d**omain but published in a **N**on-LIS **j**ournal (LD-NJ);
- Articles studied a **N**on-LIS **d**omain but published in a **L**IS **j**ournal (ND-LJ);
- Articles studied a **N**on-LIS **d**omain and published in a **N**on-LIS **j**ournal (ND-NJ).

Citations of Application Articles

To track the impact of application articles, the information of 4,178 articles that cited the application articles were downloaded from Web of Science Core Collection. We coded each article as a LIS or non-LIS article according to if it was published in a LIS journal.

RESULTS

		Journal Domain		**Total**
		LIS	**Non-LIS**	
Analyzed Domain	**LIS**	57 (11.5%)	6 (1.2%)	63 (12.7%)
	Non-LIS	115 (23.2%)	318 (64.1%)	433 (87.3%)
Total		172 (34.7%)	324 (65.3%)	496 (100%)

Table 1. An overview of application articles.

A summary of application articles is shown in Table 1. We excluded the category of LD-NJ in our study, because it only covers six articles in our dataset, most of which are in the fields of health informatics and knowledge management. The rest of 490 application articles grouped in three categories (LD-LJ, ND-LJ, and ND-NJ) were used hereafter.

Evolution of Science Mapping Applications

Figure 2. An evolution of SMT application articles in the three categories. The color is used to show the paper category and the shade of each color is used to depict tools in each category.

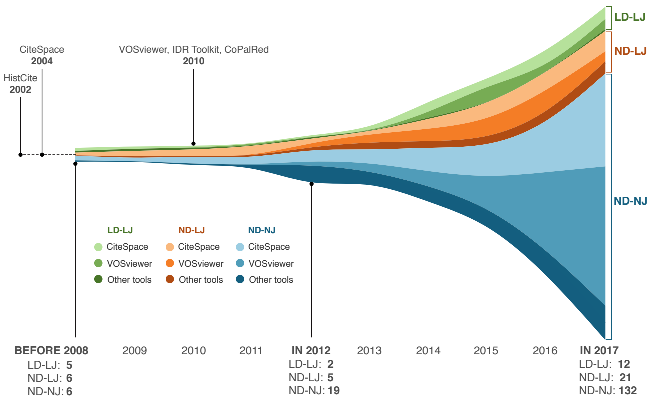


Figure 2 shows the number of three kinds of application articles from 2008 to 2017. The majority of articles in our sample (481 out of 490) were publish during this period. Most of application articles were published in LIS journals (11 out of 17) that include 5 articles studied LIS domains and 6 articles studied non-LIS domains before 2008. At that time, LIS researchers contributed most application articles, which is reasonable as science mapping itself is a LIS domain. This situation greatly changed after 2012 when there was a burst in the number of articles published in non-LIS journals where 19 articles were published. The burst was followed by a rapidly increasing trend in publishing SMT application articles in non-LIS journals, which demonstrates the extended scientific impact of science mapping as a research method. It should be noted that various tools, including CiteSpace, IDR Toolkit, VOSViewer, and CoPalRed, have been used by in the articles published in non-LIS journals.

The number of articles published in LIS journals has also been increasing, but not as fast as those in non-LIS journals in recent years. SMTs were developed for analyzing various scientific domains. We believe publishing a SMT application articles in journals focusing on a domain can be beneficial for disseminating the knowledge created by using the SMT, since the article can get exposure to scientists in the domain.

Besides visualizing the evolving patterns of the three categories, we also addressed the articles that applied two most widely used SMTs, CiteSpace (197 articles, 40.2%) and VOSViewer (192 articles, 39.2%) respectively. The rest of SMTs following CiteSpace and VOSViewer are HistCite (38, 7.8%), IDR Toolkit (25, 5.1%), and CoPalRed (13, 2.7%). CiteSpace and VOSViewer were the leading tools in all the three kinds of application articles. In comparison with articles studied non-LIS domains (ND-LJ and ND-NJ), articles studied LIS domains (LD-LJ) tended to more commonly use CiteSpace and VOSViewer.

Citation Analysis

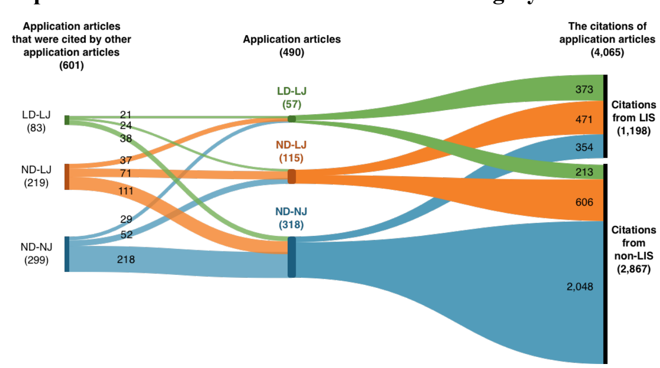
Table 2 shows an overview of the citations of applications articles. The articles in each category have a similar citation pattern in terms of their mean value and standard deviation. However, the citations of articles from the three categories are not comparable, as they were published in different disciplines and different years.

		Journal Domain		Total
		LIS	Non-LIS	
Analyzed Domain	LIS	585 (*M*=13.6, *SD*=20.7)	-	585 (*M*=13.6, *SD*=20.7)
	Non-LIS	1,077 (*M*=11.2, *SD*=14.1)	2,402 (*M*=9.6, *SD*=17.1)	3,479 (*M*=10, *SD*=16.4)
Total		1662 (*M*=12.0, *SD*=16.5)	2,402 (*M*=9.6, *SD*=17.1)	4,064 (*M*=10.6, *SD*=17.0)

Table 2. An overview of articles cited the applications articles. *M*= mean value, *SD*=standard deviation.

We focus on analyzing the citation flow of three kinds of application articles to unfold their roles in advancing the application of SMT in the scientific community. We mapped out the citation between the three kinds of application articles and the citations of application articles in Figure 3.

Figure 3. Citation flow of application articles. (#) represents the number of articles in a category. The numbers in the shades represent the citation/cited times toward a category.



On the left part of Figure 3, the composition of references of application articles in every category exist in all three categories. Articles published in non-LIS journals (ND-NJ) are less likely to be cited by articles in LIS journals. However, a considerable number of articles in LIS journal were cited by articles in non-LIS journals, especially articles studied non-LIS domains but published in LIS journals (ND-

LJ). For instance, 115 articles in ND-LJ were cited 111 times by 318 articles in ND-NJ.

The citation patterns of application articles on the right part of Figure 3 are clear. The citations of articles in LD-LJ were from articles published in LIS journals, but also have some citations from non-LIS journals. Although articles in ND-LJ were published in LIS journal, they attracted even more attention from non-LIS communities who cited ND-LJ for 606 times which is larger than the citation from LIS journal (471 times). Relevantly, articles in ND-LJ which focused on non-LIS domains but published in non-LIS journals get much fewer citations from the LIS community.

Articles in ND-LJ, it seems, played an important role in advancing the wide application of SMT. For the LIS community, the articles validate the effectiveness of SMT for analyzing various domains. They also provided practical guidelines and successful showcases for non-LIS communities.

CONCLUSION AND FUTURE WORK

In this study, we investigated the application of fourteen SMTs by visualizing 490 articles that applied SMTs to analyze scientific domains. We examined the impact of these application articles by analyzing their citations. Both journal domains and analyzed domains were addressed in our study.

A few observations concerning the roleplay of scientific mapping method and tools in the scholarly communication emerge in our approach. SMTs have been widely applied in various domains out of Library and Information Science. We have observed a rapid increasing of SMT application after 2012. CiteSpace and VOSViewer are the leading tools in both LIS and non-LIS domains. We have discovered the important role of articles, which studied non-LIS domains but published in LIS journals, in advancing the application of SMTs.

In the future, we will investigate the effects of collaboration on the application of SMTs in non-LIS domains, especially the collaboration between domain experts and scientometrics researchers. Another direction is to investigate the citation context of application articles to improve our understanding of the scientific impact of science mapping.

ACKNOWLEDGMENTS

We wish to thank the support of the National Science Foundation (Grant Number: 1633286) and the National Social Science of China (Grant number: 17CTQ025).

REFERENCES

Callon, M., Courtial, J. P., & Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. *Scientometrics*, *22*(1), 155–205.

Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, *57*(3), 359–377.

Chen, C. (2017). Science mapping: a systematic review of the literature. *Journal of Data and Information Science*, *2*(2), 1–40.

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology*, *62*(7),1382-1402.

Herman, I., Melancon, G., & Marshall, M. S. (2000). Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics*, *6*(1), 1–21.

Hey, A. J. G., Tansley, S., Tolle, K. M., & others. (2009). *The fourth paradigm: data-intensive scientific discovery* (Vol. 1). Microsoft research Redmond, WA.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, *102*(46), 16569–16572.

Kessler, M. (1963). Bibliographic Coupling Between Scientific Papers. *American Documentation*, *14*(1), 10–25.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard university press.

Latour, B., & Woolgar, S. (1979). *Laboratory of life: The construction of scientific facts*. Princenton University.

Pan, X., Yan, E., Cui, M., & Hua, W. (2018). Examining the usage, citation, and diffusion patterns of bibliometric mapping software: A comparative study of three tools. *Journal of Informetrics*, *12*(2), 481–493.

Shneiderman, B. (1996). The eyes have it: a task by data type taxonomy for informatio nvisualizations. *Proceedings 1996 IEEE Symposium on Visual Languages*, 336–343.

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.

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology,translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, *19*(3), 387–420.

White, H. D., & McCain, K. W. (1998). Visualizing a discipline: An author co-citation analysis of information science, 1972--1995. *Journal of the American Society for Information Science*, *49*(4), 327–355.