The Power to Influence: An Informetric Analysis of the Works of Hope Olson†

Dietmar Wolfram

University of Wisconsin-Milwaukee, School of Information Studies,
P.O. Box 414, Milwaukee, WI, U.S.A. 53201, USA, <dwolfram@uwm.edu>

Dietmar Wolfram is Professor in the iSchool at the University of Wisconsin-Milwaukee. His primary research interests include scholarly communication, applied informetrics, information retrieval (IR) and the intersection of IR and informetrics. He is currently Vice-President/President-Elect of the Association for Library and Information Science Education.

Wolfram, Dietmar. 2016. "The Power to Influence: An Informetric Analysis of the Works of Hope Olson." *Knowledge Organization* 43: 331-337. 12 references.

Abstract: This paper examines the influence of the works of Hope A. Olson by conducting an ego-centric informetric analysis of her published works. Publication and citation data were collected from Google Scholar and the Thomson Reuters Web of Science. Classic informetrics techniques were applied to the datasets including co-authorship analysis, citer analysis, citation and co-citation analysis and text-based analysis. Co-citation and text-based data were analyzed and visualized using VOSviewer and CiteSpace, respectively. The analysis of her citation identity reveals how Dr. Olson situates her own research within the knowledge landscape while the analysis of her citation image reveals how others have situated her work in relation to the authors with whom she has been co-cited. This reflection of Dr. Olson's research contributions reveals the influence of her scholarship not only on knowledge organization but other areas of library and information science and allied disciplines.

Received: 25 January 2016; Revised: 27 February 2016; Accepted: 3 March 2016

Keywords: informetric analysis, textual analysis, citations, Hope A. Olson

† The author would like to thank Sukjin You for assisting with portions of the data analysis.

1.0 Introduction

I first met Hope Olson at a Canadian Association for Information Science/Association Canadienne de les Sciences de l'Information conference held at the Université de Sherbrooke in 1999. We had begun our careers in academe at about the same time but had not crossed paths until then. Her research milieu was knowledge organization from a critical, humanities-based perspective, whereas mine was focused on information retrieval and informetrics employing quantitative, positivist approaches. Over the next several years we had the opportunity to become better acquainted at subsequent CAIS/ACSI meetings and to gain an appreciation for each other's research. I didn't know at the time that within a few years we would become colleagues at the University of Wisconsin-Milwaukee (UWM). During these first few years at UWM, we had many discussions about our respective research interests and noted commonalties in our research agendas and how the methodological approaches we each used could complement one another. Hope had been conducting an in-class exercise in her information organization class that demonstrated the inconstancy of inter-indexer consistency. She noted that there appeared to be a strong power law relationship in the overlap of indexing terms for larger groups of indexers. Several collaborative papers and presentations on inter-indexer consistency and tagging resulted from this discussion over the next four years (Olson and Wolfram 2006, 2008; Wolfram, Olson and Bloom 2009).

In honor of Hope's contributions to scholarship, and most notably to knowledge organization, this study provides an informetric analysis of her works by applying both traditional and newer approaches to examine the impact of her scholarship. Although metrics research has fre-

quently been used to understand regularities in the way information is produced and used at a disciplinary or subdisciplinary level, it also lends itself to detailed examination of individuals and their influence on the academic landscape. The title of this paper borrows from Dr. Olson's most cited work: The *Power to Name: Locating the Limits of Subject Representation in Libraries* (Olson 2002). In this book, she outlines how early classification pioneers established conventions that have biased or influenced representation. In a sense, what researchers do is similar. A scholar's products, in the form of publications, may influence both contemporary and future researchers. The impact of scholars and their scholarly products may be assessed in different ways. Over the past half-century, a number of techniques and metrics have been developed to assess scholarly influence.

At the heart of metric investigations is a set of measures based on data associated with scholarly products. These include citation-based measures, those based on collaboration and, increasingly, the textual content of an author's *oeuvre* or body of work. Citation analysis, in particular, has been used to assess the influence of an individual, publication or journal based on the references they provide in their own work, the citations their works receive and the works and people with whom they are co-cited (i.e., cited together in the same work). The relationship between citation analysis and knowledge organization has been studied by Hjørland (2013), who noted that mapped relationships based on co-citation analysis may be used to represent social knowledge organization systems.

White (2000) proposed the use of ego-centric citation analysis to gain insight in the influence of researchers by examining not just the citations received but the origin of the citations, namely the citers. He introduced two concepts related to citer analysis. A "citation identity" consists of all the authors a citing author cites. "Citation image makers" are those authors who cite an author of interest. White also delineated concepts related to the co-citation of authors and introduced the idea of a "citation image," which consists of all authors with whom an author of interest has been co-cited. To demonstrate these concepts, White (2001) compiled the citer patterns for eight prominent information scientists. White's call for more citer-based research has been answered by several researchers. Cronin and Shaw (2002), for example, explored the citation identities of three information science scholars. Included in their analysis was the total number of citers along with more traditional citation-based measures. Similarly, Bar-Ilan (2006) conducted an ego-centric analysis of computer scientist Michael O. Rabin that incorporated White's ideas as well as co-authorship data. The present contribution takes a similar approach by applying citation, co-author and language-based analysis of Hope Olson's body of work.

2.0 Method

Records of Dr. Olson's publications were initially identified in Thomson Reuters Web of Science (WoS) Core Collection using 'Olson H*' as the author in early September 2015 for the full length of UW-Milwaukee's WoS subscription period (1987-2015). The returned items were processed to remove publications not authored by Hope along with minor works such as book reviews or letters. This set of 34 publications, spanning 1995 to 2012, represented the subset of Dr. Olson's *oeuvre* indexed by WoS. These records served as the core dataset for the citation analyses conducted. WoS allows searchers to download the full bibliographic records of retrieved items, which includes bibliographic details of the sources referenced in each document, representing Hope's citation identity. Using the retrieved publication set, the Create Citation Report feature of WoS was then selected to provide citation data regarding Hope's publications. By then clicking on the Citing Articles link on the citation report page, a list of all publications that have cited the list of Hope's papers is produced. Selecting the Analyze Results feature, followed by Authors and then selecting the desired scope produces a list of all citers (citation image makers). By downloading the full records along with the references of these records, authors who have been co-cited with Dr. Olson may be identified (citation image). Self-citations (i.e., citing records authored by Hope) were removed from this dataset so that the analysis reflected only the influence of citing authors other than Hope.

To analyze Hope's citation data, the WoS citing and citation records were imported into two scientific visualization analysis software packages. VOSviewer (van Eck & Waltman 2009) permits the visualization of bibliographic networks as nodes and links based on extracted data such as authors and their co-citations. The influence of nodes (authors, journals, papers) and their relationships are visually represented as proximities and link strengths based on co-occurrences. Related nodes may be clustered into groups identifying further affinities. VOSviewer was used to visualize Hope's citation identity and citation image. Although VOSviewer also possesses text analysis capabilities, another visualization package, CiteSpace (Chen 2006), was used because of its ability to identify relationships based on phrases and not just explicit terms. Like VOSviewer, CiteSpace uses citation and textual data from bibliographic sources such as WoS to explore relationships among bibliographic entities of interest as networks of nodes and links. It can also process citation and co-citation networks. CiteSpace also has text visualization abilities that incorporate natural language processing techniques on noun phrases or keywords. The software allows data to be viewed through selected time slices. For the present analy-

sis, a single time slice was used for the text appearing in sources published by Hope and the identified publications that cite Hope's work due to the relatively small dataset.

It should be noted that not all of Dr. Olson's research contributions are indexed in Web of Science. This is one current limitation associated with relying on a citation database such as WoS or Elsevier's Scopus (which was not available for the study). Although providing a rich set of data for those sources that are indexed, other sources such as monographs and conference proceedings, which are common dissemination outlets in LIS, may not be included. For this reason, citation data were also collected from Google Scholar (GS) using Publish or Perish (http://www.harzing.com/pop.htm), which casts a broader net than WoS and Scopus, and may result in more publications being reported and higher citation counts. Despite casting a broader net, neither Google Scholar nor Publish or Perish provides direct access to the same rich array of citation-based data or the same range of search features. As a result, only citation counts for the GS data are reported. The counts are higher than those reported by WoS for those publications WoS indexed.

3.0 Results

3.1 Summary data

Google Scholar lists 99 publications (monographs, journal articles, conference proceedings, editorials and other documents) with over 1,100 citations and an h-index of 17. An author's h-index (Hirsch 2005), *y,* is a measure of influence where an author has published *y* items that have each been cited at least *y* times. Based on her listed *oeuvre* in GS, she has co-authored with 34 students, academics and professional colleagues (Table 1).

3.2 Citation identity

A researcher's citation identity reflects the influences of a researcher's work. It also reveals how authors situate themselves within a disciplinary landscape based on their citing behavior. The resulting VOSviewer co-citation visualization for Hope's citation identity for the 34 authors cited at least three times appears in Figure 1. Note that some authors with close proximity to others may have their labels hidden by the software. The size of the nodes indicates the relative importance of the respective author based on citation totals. The most frequently cited authors in Hope's *oeuvre* (other than her own work) are: Dewey, M; Cutter, CA; Aristotle; Cleverdon, CW; Foskett, AC; Markey, K; OCLC; and Swanson, DR. Most of these authors reflect Dr. Olson's interests in cataloguing and classification or are pioneering researchers in information science. VOSviewer will automatically cluster nodes on the map with clusters of citing behavior represented by different colors. The clusters were limited to membership of at least six authors given the relatively small dataset from which cluster members were determined. Also, the smaller number of total clusters provides a larger number of members per cluster to assist in the interpretation of the agglomerations. The clusters roughly correspond to three primary research areas of Hope's interests. The green cluster with members on the upper and left part of the map represents aspects of cataloging and classification applications and education. It's not surprising that the Book Industry Study Group and OCLC are not closely aligned to specific authors, so they are situated more distantly from the authors. The blue cluster on the right side represents authors who investigate subject access, indexing and information retrieval, corresponding to Hope's interests in areas such as subject access and inter-indexer consistency. The red cluster, for which several labels do not appear due to their close proximity, primarily address Dr. Olson's interests: philosophy and feminist perspectives in knowledge organization as well as some pioneers in information science research.

3.3 Citation image makers and citation image

Dr. Olson's citation image makers—that is, those who have cited Hope's work as recorded by WoS—comprise 137 authors. These include authors who have published in disciplinary areas such as library and information science (LIS), communication, education, the natural sciences and the humanities. The citation image makers help to situate Dr. Olson's work by creating her citation image, which identifies the authors with whom Hope has been co-cited.

The references in the citation image makers' papers provided the co-citation data that represent her citation image. The raw dataset was cleaned to remove selected records. Citations to sources marked Anonymous were excluded. Similarly, one author, Makarov MN, was excluded due to the far proximity of the author from the main body of the map. Co-citations arising from Hope's own work were removed to eliminate the influence of self-citations. The top 68 authors, consisting of those authors for which there were at least seven citations, were included in the VOSviewer co-citation analysis mapping (Figure 2). Clusters were set to a minimum size of four authors.

The five authors with whom Hope has been cited most frequently are: Hjørland, B; Floridi, B, Foucault, M; Beghtol, C; and Frohmann, B. This list contains philosophers and philosophy-influenced researchers, reflecting how much as Dr. Olson's research has also been informed by these areas. Although frequently co-cited with Dr. Olson, the locations of authors in the map are influenced by every author in the analysis. The five clusters reflect broad areas

Citations	Author	Title	Year
142	HA Olson	The power to name: locating the limits of subject representation in libraries	2002
109	HA Olson	Mapping beyond Dewey's boundaries: Constructing classificatory space for marginalized knowledge domains	1998
81	HA Olson	The power to name: Representation in library catalogs	2001
64	HA Olson, JJ Boll	Subject analysis in online catalogs	2001
60	H Olson	Quantitative" versus" qualitative research: The wrong question	1995
54	HA Olson	How we construct subjects: A feminist analysis	2007
54	HA Olson, R Schlegl	Standardization, objectivity, and user focus: A meta-analysis of subject access critiques	2001
47	HA Olson	Sameness and Difference	2011
44	HA Olson	Difference, culture and change: The untapped potential of LCSH	2000
43	HA Olson	Exclusivity, teleology and hierarchy: Our Aristotelean legacy	1999
29	LM Given, HA Olson	Knowledge organization in research: A conceptual model for organizing data	2003
27	HA Olson	The feminist and the emperor's new clothes: Feminist deconstruction as a critical methodology for library and information studies	1997
25	HA Olson	The ubiquitous hierarchy: An army to overcome the threat of a mob	2004
19	HA Olson, D Wolfram	Syntagmatic relationships and indexing consistency on a larger scale	2008
18	A Kublik, V Clevette, D Ward, HA Olson	Adapting dominant classifications to particular contexts	2003
18	HA Olson	Patriarchal structures of subject access and subversive techniques for change	2001
17	HA Olson, DB Ward	Feminist locales in Dewey's landscape: Mapping a marginalized knowledge domain	1997
17	HA Olson, D Wolfram	Indexing consistency and its implications for information architecture: a pilot study	2006
17	HA Olson	Between control and chaos: an ethical perspective on authority control	1996
16	H Olson	Subject access to women's studies materials	1991
16	D Wolfram, HA Olson	A method for comparing large scale inter-indexer consistency using IR modeling	2013
15	HA Olson	Dewey thinks therefore he is: The epistemic stance of Dewey and DDC	1996
14	HA Olson	The power to name: Marginalizations and exclusions of subject representation in library catalogues	1996
13	B Mehra, HA Olson, S Ahmad	Integrating diversity across the LIS curriculum: An exploratory study of instructors' perceptions and practices online	2011
12	D Wolfram, HA Olson, R Bloom	Measuring consistency for multiple taggers using vector space modeling	2009
12	HA Olson	Thinking professionals: teaching critical cataloguing	1997
12	HA Olson	Universal models: a history of the organization of knowledge	1994
11	HA Olson	Cultural discourses of classification: Indigenous alternatives to the tradition of Aristotle, Durkheim and Foucault	1999
10	HA Olson, DB Ward	Charting a journey across knowledge domains: Feminism in the Dewey Decimal Classification	1998
9	HL Lee, HA Olson	Hierarchical navigation: An exploration of Yahoo! directories	2005

Table 1. Citation summary of Hope Olson's most cited works (Google Scholar)

of LIS and allied subject areas, demonstrating the relevance of Hope's work across LIS and other fields. The red cluster in which Hope features prominently includes authors whose works address information organization and retrieval. The green cluster with Birger Hjørland includes knowledge organization researchers who are known for humanistic research approaches. The blue cluster includes researchers more broadly in LIS or other humanities. The more scattered magenta cluster includes researchers who have contributed to historical or foundational aspects of the field. The yellow cluster consists of researchers who use more quantitative or informetrics-based approaches in LIS. The range of areas represented indicates the Dr. Olson's work has application across the breadth of methodological approaches and topical areas of LIS.

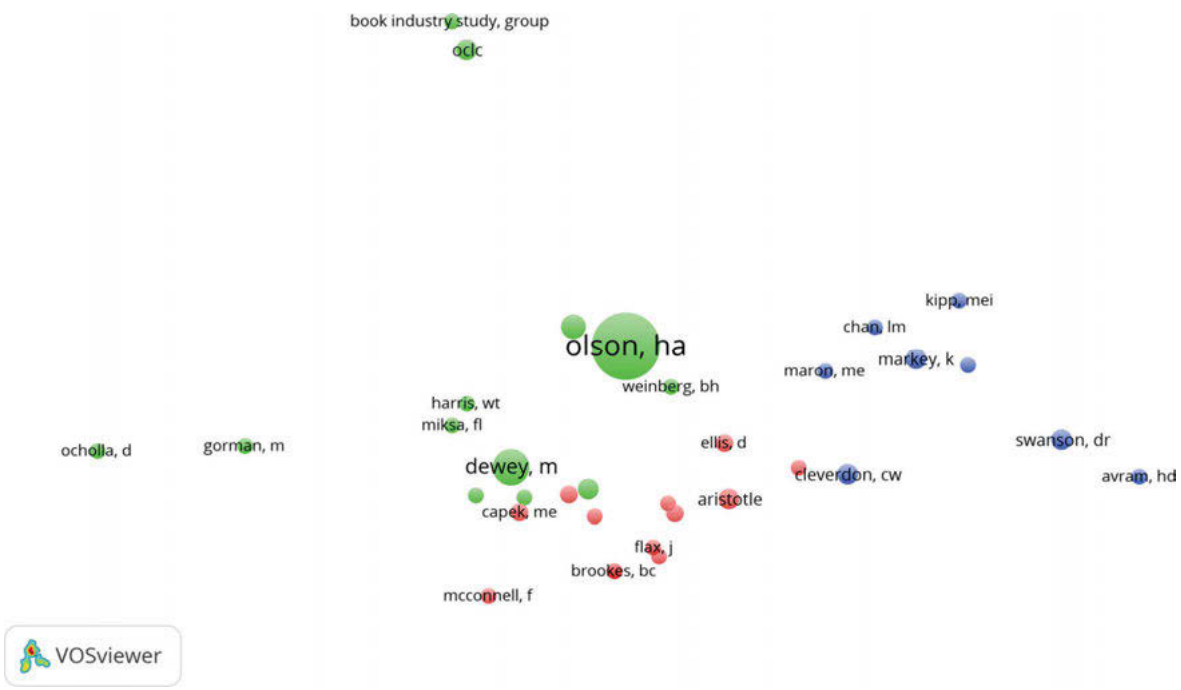

Figure 1. Citation identity map for Hope Olson

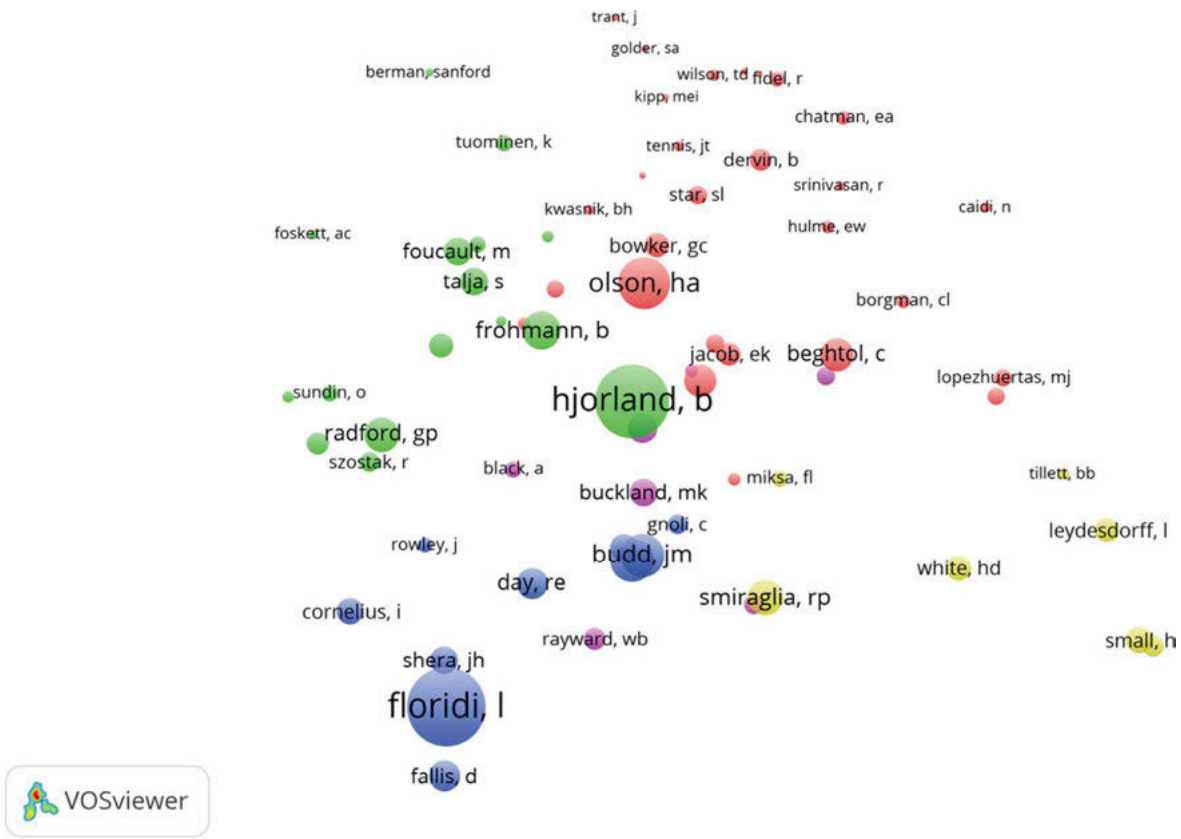

Figure 2. Citation image map for Hope Olson

3.4 Text-based themes

CiteSpace performs selected textual analyses where noun phrases associated with publication titles and abstracts may be analyzed. A network of the 30 most frequent noun phrases was constructed for Dr. Olson's own publications based on text appearing in the titles and abstracts of the retrieved records. Unlike a simple word cloud, which only lists terms based on frequency, the CiteSpace textual analysis also provides a network analytic context for the terms or phrases. Figure 3 summarizes the most prominent noun phrases present in Dr. Olson's work. One can see the significance of knowledge organization-related phrases, in particular those related to classification and allied areas that have influenced her research such as information retrieval and feminist deconstruction. The thickness of the lines indicates the strength between the phrases. Because the analysis only focuses on records indexed by WoS, the representation of research areas published as part of conference proceedings or monographs may not be as strongly reflected in the map due to the more limited coverage.

When the same analysis is applied to the title and abstract fields associated with Hope's citation image makers for the top 50 noun phrases, what becomes evident is that her contributions have influenced others' work in not only knowledge organization but other areas of library and information science as well. For instance, themes of library practice, research methods and information science are represented in the map (Figure 4).

4.0 Conclusion

The use of informetric techniques that rely on publication, citation and textual data of authors can shed light not only on larger entities of analysis such as disciplines but also can be revealing of the contributions and impact of individuals. Metric techniques may also be used to represent social knowledge organization systems (Hjørland 2013). This study has investigated the influence of one noted scholar in knowledge organization using these techniques. Hope Olson has built an international reputation as a scholar and educator over a professional and academic career that has spanned several decades. Of particular note are her contributions to subject representation and classification, inter-indexer consistency, feminist perspectives and research methods. The co-citation analyses, both from the citation identity and citation image perspectives, reveal a scholar who not only is influenced by researchers in a range of areas of LIS and other fields but whose research is also cited in many areas of LIS and other disciplines. The textual analysis revealed equally broad influences based on the noun phrases present in Hope's work and those used by authors who have cited her work. Dr. Olson's contributions demonstrate that one individual can indeed influence (or inspire) researchers within and across disciplines.

References

Bar-Ilan, Judit. 2006. "An Ego-centric Citation Analysis of the Works of Michael O. Rabin Based on Multiple Citation Indexes." *Information Processing & Management* 42:1553-66.

Chen, Chaomei. 2006. "CiteSpace II: Detecting and Visualizing Emerging Trends and Transient Patterns in Scientific Literature." *Journal of the American Society for Information Science and Technology* 57:359-77.

Cronin, Blaise and Debora Shaw. 2002. "Identity-creators and Image-makers: Using Citation Analysis and Thick Description to Put Authors in their Place." *Scientometrics* 54:31-49.

Hirsch, J. E. 2005. "An Index to Quantify an Individual's Scientific Research Output." *Proceedings of the National academy of Sciences of the United States of America* 102:165 69-72.

Hjørland, Birger. 2013. "Citation Analysis: A Social and Dynamic Approach to Knowledge Organization." *Information Processing and Management* 49:1313-25.

Olson, Hope A. 2002. *The Power to Name: Locating the Limits of Subject Representation in Libraries*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Olson, Hope A. and Dietmar Wolfram. 2006. "Indexing Consistency and its Implications for Information Architecture: A Pilot Study." *IA Summit*. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.105.6370&rep=rep1&type=pdf

Olson, Hope A. and Dietmar Wolfram. 2008. "Syntagmatic Relationships and Indexing Consistency on a Larger Scale." *Journal of Documentation* 64:602-15.

van Eck, Nees and Ludo Waltman. 2009. "Software Survey: VOSviewer, a Computer Program for Bibliometric Mapping." *Scientometrics* 84:523-38.

White, Howard D. 2000. "Toward Ego-centered Citation Analysis." In *The Web of Knowledge*, ed. Blaise Cronin and Helen Barsky Atkins. ASIS Monograph Series. Medford, NJ: Information Today, 475-96.

White, Howard D. 2001. "Authors as Citers Over Time." *Journal of the American Society for Information Science and Technology* 52:87-108.

Wolfram, Dietmar, Hope A. Olson and Raina Bloom. 2009. "Measuring Consistency for Multiple Taggers Using Vector Space Modeling." *Journal of the American Society for Information Science and Technology* 60:1550-71.

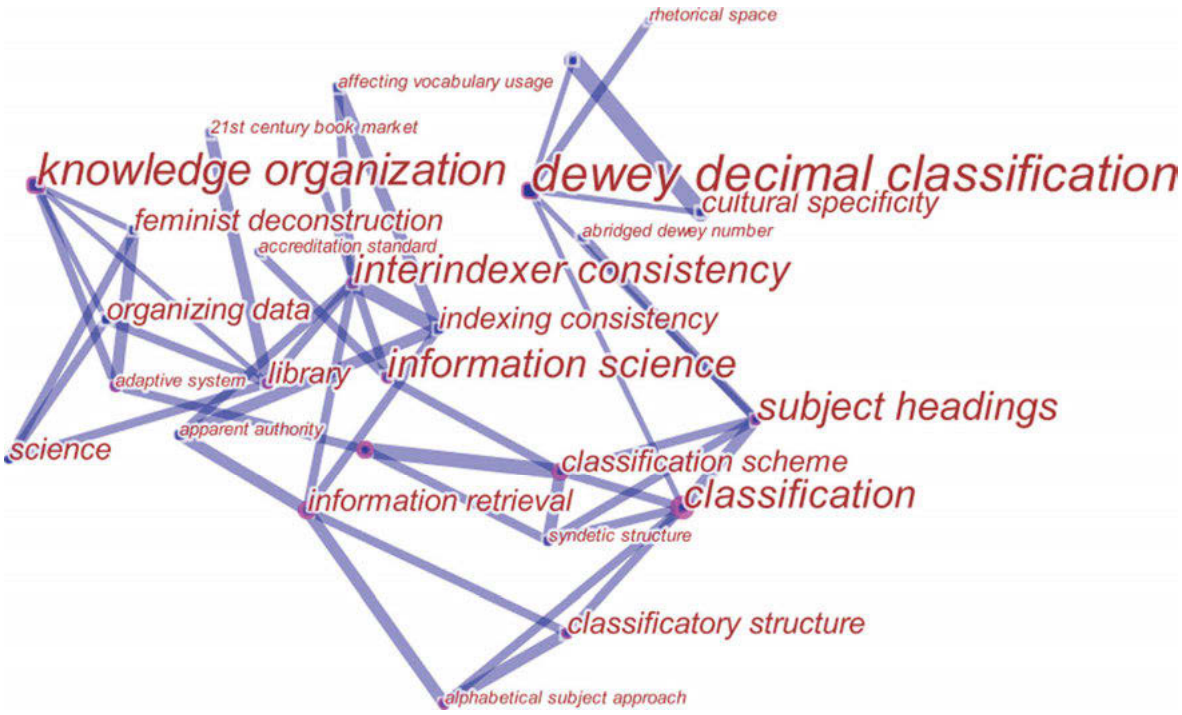

Figure 3. Relationships among identified phrases and concepts in Hope Olson's work

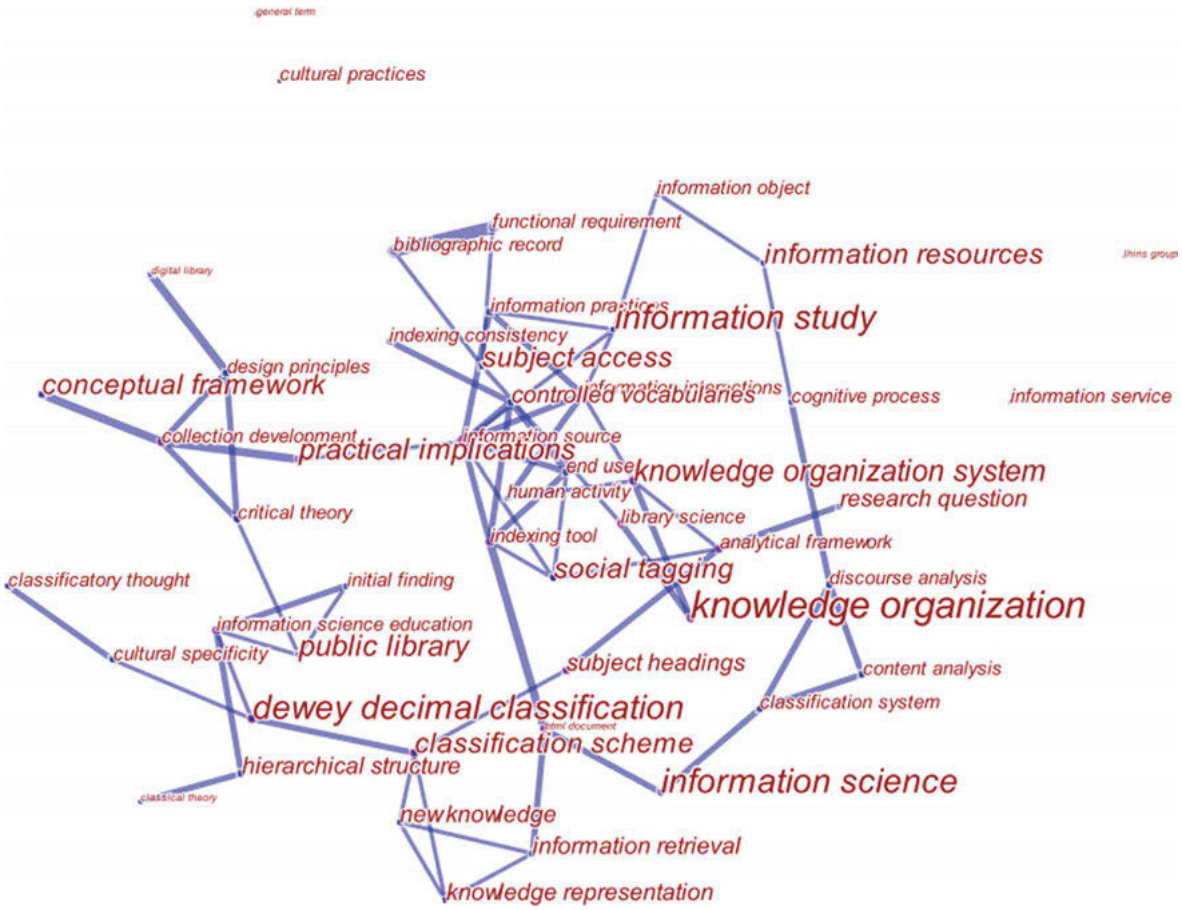

Figure 4. Relationships among identified phrases and concepts in articles citing Hope Olson's work