Visualizing and Comparing Four Facets of Scholarly Communication: Producers, Artifacts, Concepts, and Gatekeepers[1]

Chaoqun Ni, Cassidy R. Sugimoto, Blaise Cronin
School of Library and Information Science
Indiana University
1320 East Tenth Street
Bloomington, IN 47401
{chni, sugimoto, bcronin}@indiana.edu

ABSTRACT

This paper extends Borgman's (1989) three-facet framework (artifacts, producers, concepts) for bibliometric analyses of scholarly communication by adding a fourth: gatekeepers. The four-facet framework was applied to the field of Library and Information Science to test for variations in the networks produced using operationalizations of each of these four facets independently. Fifty-eight journals from the Information Science & Library Science category in the 2008 Journal Citation Report were studied and the network proximity of these journals based on Venue-Author-Coupling (producer), journal co-citation analysis (artifact), topic analysis (concept) and interlocking editorial board membership (gatekeeper) was measured. The resulting networks were examined for potential correlation using the Quadratic Assignment Procedure. The results indicate some consensus regarding core journals, but significant differences among some networks. Holistic measures of scholarly communication that take multiple facets into account are proposed. This work is relevant in an assessment-conscious and metrics-driven age.

KEYWORDS

Producers, Artifacts, Concepts, Gatekeepers, Scholarly Communication, Bibliometrics

INTRODUCTION AND BACKGROUND

Bibliometric research is a well-established quantitative approach for the investigation of formal channels of scholarly communication. Borgman (1989) proposed three main facets by which scholarly communication can be examined: producers, artifacts, and concepts. According to her definition, producers are individual authors or aggregations of authors, artifacts are publication units or aggregations of such units, and concepts are expressed by authors' word choices and the indexing terms assigned by others. Paisely (1990) suggested that the matrix could be expanded.

[1] This paper is extended from a poster presented at the 2011 *Annual Meeting of American Society for Information Science and Technology* (Ni & Sugimoto, 2011).

The bibliometric literature is replete with studies that focus on each of these facets. There has been more than a century of research at the artifact level (Gross and Gross 1927; Cole and Eales 1917; Egghe and Rousseau 1990; Sengupta 1992) with large-scale analysis made possible by the introduction of the *Science Citation Index* (Garfield 1955). Citation analyses of artifacts have been conducted using bibliographic coupling and co-citation analysis (Kessler 1963a, 1963b, 1965; Small 1973; McCain 1991; Jarneving 2007; Sugimoto et al. 2008; Small and Koenig 1977).

Producers have been studied using citations and references as units of coupling (Zhao and Strotmann 2008b, 2008a; White and Griffith 1981; Ding 2011). Collaboration has also been examined, operationalized as co-authorship. Beaver and Rosen led the way in studies of scientific collaboration (Beaver and Rosen 1978, 1979a, 1979b). Subsequent research (e.g. Liu et al. 2005) focused on various levels of aggregation (e.g., institutions (Thijs and Glanzel 2010; Van Rijnsoever et al. 2008), fields (Fox 2008), and countries (Wagner 2005; Luukkonen et al. 1992; He 2009)). Journals have also been clustered according to shared author profiles using the Venue-Author-Coupling (VAC) method (Ni et al. to appear).

Concept analysis is based on the assumption that words are indicative of concepts and describe the content of the article of which they are a part (Leydesdorff 1997). Single word frequencies and co-occurrences of words in texts have been analyzed (Callon et al. 1983; Rip and Courtial 1984; Milojević et al. 2011; Teichert et al. 2011). Topic analysis has been the focus of much recent work, with approaches such as Latent Dirichlet Allocation (Blei et al. 2003) being used to reveal major topic distributions over large corpora of scholarly texts (Sugimoto et al. 2011; Yan et al. 2012).

The journal editorial board is another important entity in the scholarly communication process. Scientific journals first appeared in the 17th century and remain the dominant scholarly publication genre in many fields (Braun 2005). Most journals of stature maintain an editorial board, in addition to a corps of ad hoc reviewers. These populations are not, however, mutually exclusive. Many editorial board members are involved in reviewing (Cronin 2009; Glogoff 1988). There are two principal objectives for an editorial board: to establish and uphold quality criteria and to chart and maintain a course for the journal. In this way, editorial board members "set the scientific standards of a discipline" (Bedeian et al. 2009, p. 212) and "protect and warrant in the future the social and intellectual integrity of science" (Braun 2005, p. 96). Editorial boards are also an indicator of a journal's prestige, signaling a standard for prospective authors and readers.

Barzilai-Nahon (2009) defined "editorial gatekeeping" as a critical information filtering role, both in journalism and scientific publishing. Quantitative studies of editorial gatekeepers, however, are scarce, despite the central role played by these actors (Bedeian et al. 2009; Barzilai-Nahon 2009; Budd 2000). Studies have examined the relationship between the scientific achievements of editorial board members and the prestige of the journal with which they are associated (Braun and Bujdosó 1983; Braun and Dióspatonyi 2006; Bakker and Rigter 1985; Zsindely et al. 1982). The geopolitical effects on scholarship of international editorial boards have also been explored. Braun's (2005) work has demonstrated that there is a correlation between the country affiliation of editorial board members and the country affiliation of publishing authors. Recently, network analysis has been applied to analysis of editorial

board memberships. Two approaches, measuring journal proximity via coupling (Ni and Ying 2010) and social network analysis (Baccini and Barabesi 2011), have been utilized to describe the scholarly communication landscape from the perspective of editorial gatekeepers.

Here we examine the role played by these gatekeepers, and their relationship to the three other facets. We generate four networks using the following methods: a) topic analysis, b) co-citation analysis, c) Venue-Author-Coupling, and d) editorial board member coupling to understand how the choice of facet might affect scientometric evaluations. Specifically, we want to address the following question: Are there significant differences among the journal networks produced by these four methods? To answer this question, all producers, artifacts, concepts and gatekeepers are aggregated at the journal level to compare the proximities of a single unit based on the different facets.

METHODS AND MATERIALS

DATA COLLECTION AND PROCESSING

We used two sources of data: lists of editorial board members and Web of Knowledge. Editorial board member information was gathered from the website or from the hard copy of each of the 58 journals classified as Information Science & Library Science (IS&LS) in the Journal Citation Reports 2008[2]. All the editorial board members included are those indicated officially as of December 2009 (see Ni & Ding, 2010 for more information on this process). Web of Knowledge was used to identify all IS&LS publications, articles, and proceeding papers published between 2005 and 2009. We acknowledge that journal editorial board membership might have changed slightly over the five year time period. However, gathering year-by-year data proved extremely difficult. The laboriousness of the task notwithstanding, the data were simply not available for the majority of journals. Journal websites only provided the current editorial board member information and few journals printed editorial board information in each issue. Personal contact was made with a number of current editors who reported no documentation on editorial board membership in years preceding their role as editor. Including only those journals with complete editorial information for the entire time period would have eliminated the majority of the journals and invalidated the subsequent analyses. Therefore, a decision was made to keep all of the journals for a single year (2009), rather than only a few journals for the entire time period. The results should be interpreted in light of this limitation. However, data on the few journals that provided longitudinal editorial board membership indicated that there are not large fluctuations in the composition of editorial boards, particularly over a five year period.

APPROACHES AND TECHNIQUES

This paper generates networks of IS&LS journals using four different facets: producers, artifacts, concepts and gatekeepers.

[2] There are 61 journals categorized as Information Science & Library Science journals in the 2008 Journal Citation Report, but three of them were excluded as they are in languages other than English. The use of a non-English language would have invalidated the results of the topic modeling.

The Venue-Author-Coupling (VAC) approach is used to represent the producers in this dataset. This approach measures journal proximity based on the number of authors shared by each journal pair. The VAC approach is based on the idea that an author's choice of publication venue reflects similarity judgments—authors are likely to choose venues that are thematically or socially similar. Therefore, if an author chooses to submit his or her work to two journals, we assume that these two journals share some similarity. The degree of similarity increases with the number of authors shared by the journals (Ni et al. to appear).

Artifacts are measured by means of journal co-citation. This measure, introduced by McCain (1991), refers to the appearance of two journals in the same reference list of an article. The more frequently two journals appear in the same reference lists, the greater the similarity between the two journals. The journal co-citation approach measures journal proximity by the frequency with which each journal pair is co-cited by the same articles.

Topic modeling is used to capture concepts. The topic modeling approach measures journal proximity in terms of semantic themes. The technique adopted here, the author-conference-topic (ACT) model[3], extends the Latent Dirichlet Allocation (LDA) model by considering the author and publishing venue of the articles. LDA was developed originally as a topic modeling technique concerning the probability distribution of keywords for topics, and is particularly helpful with the "classification, novelty detection, summarization, and similarity and relevance judgment" of large-scale data (Blei et al. 203, p. 993). This modeling technique is based on the assumption that there is no inherent ordering function that exists between words and documents in a corpus, and thus each individual word can be viewed as being conditionally independent and identically distributed (Blei et al. 2003). Based on a hierarchical Bayesian model, LDA characterizes each topic by a probability distribution for each word under each topic. The technique was subsequently extended by adding more features, i.e. the publishing venue and the author (Tang, Jin and Zhang 2008). This model extends the idea of LDA by taking into account the authors and publishing venues, and estimates not only the distribution of words on topics but also the distribution of authors and venues on the topics modeled. The extended model was abbreviated as ACT (author-conference-topic), in that the original paper dealt with conference proceedings. In the current paper, journals replace conferences, but the name ACT is retained. Here, the outcome of the ACT model is the probability distribution of each author and each journal over topics, and the journal proximity is calculated using the cosine similarity of the journals. For more detail and a mathematical explanation of the ACT model, please refer to Tang et al. (2008).

The interlocking editorship approach, employed by Ni and Ding (2010), measures journal proximity based on common editorial board membership. The number of editorial board members that two journals share can be viewed as an indicator of journal similarity. This assumes that journals tend to approach scholars whose research interests reflect the journal's scope. Thus, it can be expected that if two journals have scholars in common on their editorial boards, these two journals have some degree of similarity, either cognitively or socially.

[3] The only modification to the model is the use of journals rather than conferences. Please refer to Tang et al.(2008) for details.

Network representations of journal proximities are produced using these four approaches. Statistical techniques commonly used in network analysis studies were employed. Journal proximities were expressed using cosine similarity. The journals were clustered using a hierarchical clustering technique with squared Euclidean distance and Ward's method (Steven, 2002). Each journal clustering was displayed as a network (Kamada-Kawaii layout); each node (journal) was colored according to the hierarchical clustering result with the size of a node proportional to its centrality (either degree or closeness). Additionally, a comparison of journal proximity results was conducted using the Quadratic Assignment Procedure (QAP). QAP is commonly used in social network analysis as a means of investigating correlations between two networks. Different from some common correlation tests (e.g., Pearson's *r* and Spearman) for two sets of subjects, QAP reveals the correlation between two treatments on the same set of subjects (Lawler 1963).

RESULTS

SUMMARY DATA

Table 1 provides detailed information for each of the four facets. There are 17,589 unique authors (producers) who have published in the 58 journals within the time period examined (2005-2009). Roughly 64% published once and less than 7% published more than five times in the journal set. The total number of publications (artifacts) is 13,627. The percentages of single-authored and multiple-authored papers are very close: 6,455 (47%) and 7,172 (53%), respectively. Some 16,292 concept words were identified for analysis after removing stop words. As for the gatekeepers, there are 1,785 available editorial board positions occupied by 1,561 unique board members.

Finally, it is worth noting that the number of publications that each journal has differs due to the different publishing genres and frequencies, ranging from 66 (*Annual Review of Information Science & Technology*) to 893 (*Journal of the American Society for Information Science & Technology*). This should be considered as a possible limitation when analyzing the number of unique authors per publication.

Table 1 Summary of Four Facets

Facets	Sub-items	Value
Producer	#unique authors	17,589
	#unique authors per journal	303.26
	#authors per paper	2.07
Artifact	#single-author papers	6,455
	#multi-author papers	7,172
	#papers per journal	1,269
	#total citations as of 2009	31,842
	#citations per journal 2009	549
Concept	#unique words in titles	16,292
GateKeeper	#editorial board members	1,561
	#editorial board positions	1,785

PRODUCERS

Figure 1 displays the journal proximity network obtained using the VAC approach[4]. The journals were arranged into four groups and colored according to the hierarchical clustering results. The four groups are color coded: Management Information Systems (MIS) (yellow), Information Science (IS) (blue), Library Science (LS) (green), and Specialized (red). The size of each node (journal) in the network is proportional to its degree centrality, and the width and scale of the edge are proportional to the cosine similarity between journals linked by the edge. For example, The *Journal of the American Society for Information Science and Technology* (*JASIST*) occupies the central position in the journal proximity network in terms of shared authors. It is very close to the *Annual Review of Information Science and Technology* (*ARIST*), *Information Processing & Management* (*IPM*), and *Journal of Documentation* (*JDOC*). The MIS cluster is quite separate from the other three: few of the authors publishing in the MIS journals published in the other journals. Inside the MIS cluster, some journals, e.g. *MIS Quarterly* (MIS), *Journal of Management of Information Systems* (JMIS) and *Information System Research* (ISR), are particularly close to each other in terms of shared author profiles. Specialized cluster journals are not well connected to other clusters, or internally. Journals in this cluster, e.g., *Government Information Quarterly* (*GIQ*), *Law Library Journal* (*LLJ*), *Journal of Health Communication* (*JHC*) and *International Journal of Geographical Information Science* (IJGIS), publish in specific areas of IS&LS and authors publishing in these journals may well have their primary disciplinary home elsewhere. The LS and IS clusters are more closely aligned than the other two clusters. *Library and Information Science Research* (*LISR*) and *Journal of Academic Librarianship* (JALib) seem to be bridging journals between the clusters. Inside the IS cluster, *Journal of Informetrics* (JInfMetrics), *Scientometrics* (SciMetrics) and *Research Evaluation* (ResEva) form a subcluster, likely reflecting their shared scope.

[4] In each of the following network views, cosine similarity was used as a proximity measure between journals. Some network views only display lines with values larger than 0.2 to make them more readable.

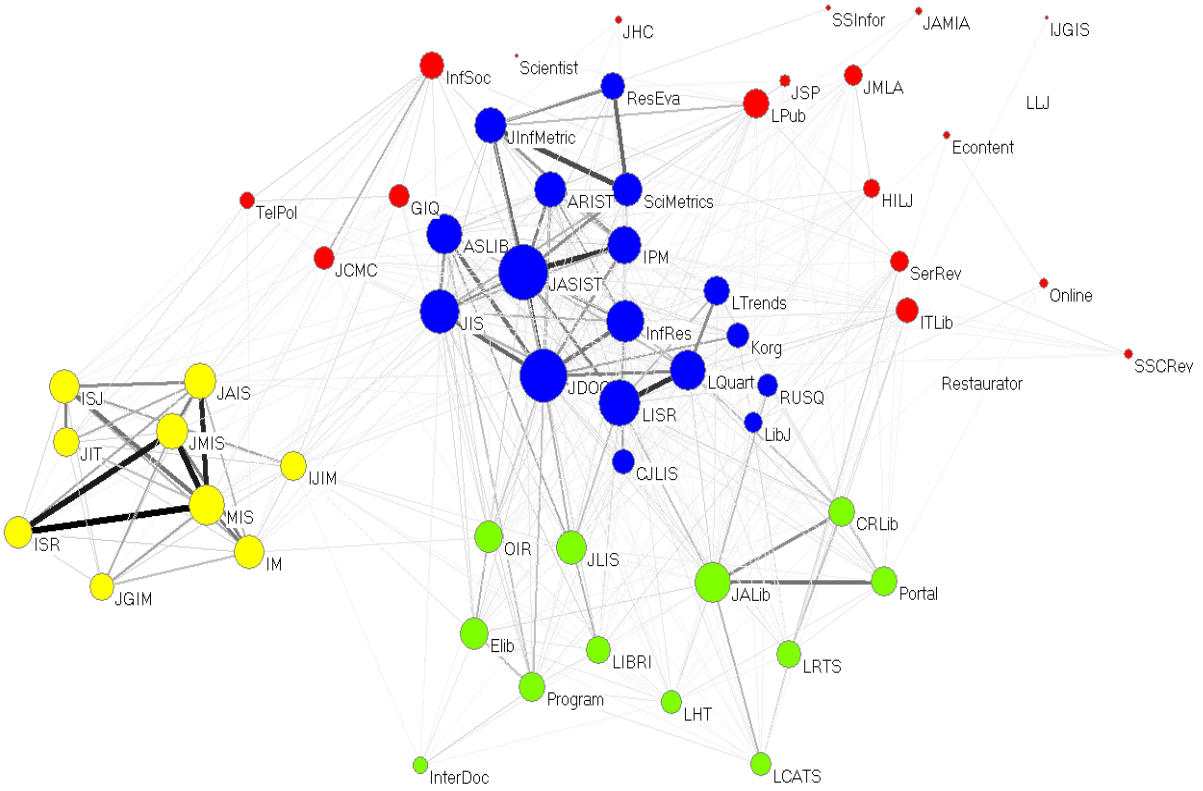

Figure 1 Network View of Journal Clusters by Producers

ARTIFACTS

The clustering technique for journal co-citation also yields four groups, displayed in Figure 2. *ARIST* and *JASIST* are in the center, while *IPM*, *SciMetrics*, *Journal of Information Science (JIS)*, and *JDOC* occupy relatively core positions in terms of degree centrality (demonstrated by the size of the node). The same nine journals in the MIS producer network are also clustered here (yellow), with the addition of *Telecommunication Policy* (*TelPol*) which serves as a bridging journal between the MIS and Specialized clusters. A large subset of the IS journals created in the producer network remains together based on the co-citation analysis (blue). However, the additional clusters differ from those of the producer network. For the remaining clusters, the journal nodes are colored by the VAC results: if a majority of journals in the cluster by journal co-citation overlaps with that of the VAC results, the cluster is colored the same. It can be observed from Figure 2 that *GIQ*, *JHC*, and *Journal of the American Medical Informatics Association (JAMIA)*, are still grouped together and colored red. However, some of the Specialized journals are more connected to the rest of the network in this arrangement and some LS journals are less connected. *ONLINE* plays a bridging role between the IS cluster and Specialized clusters.

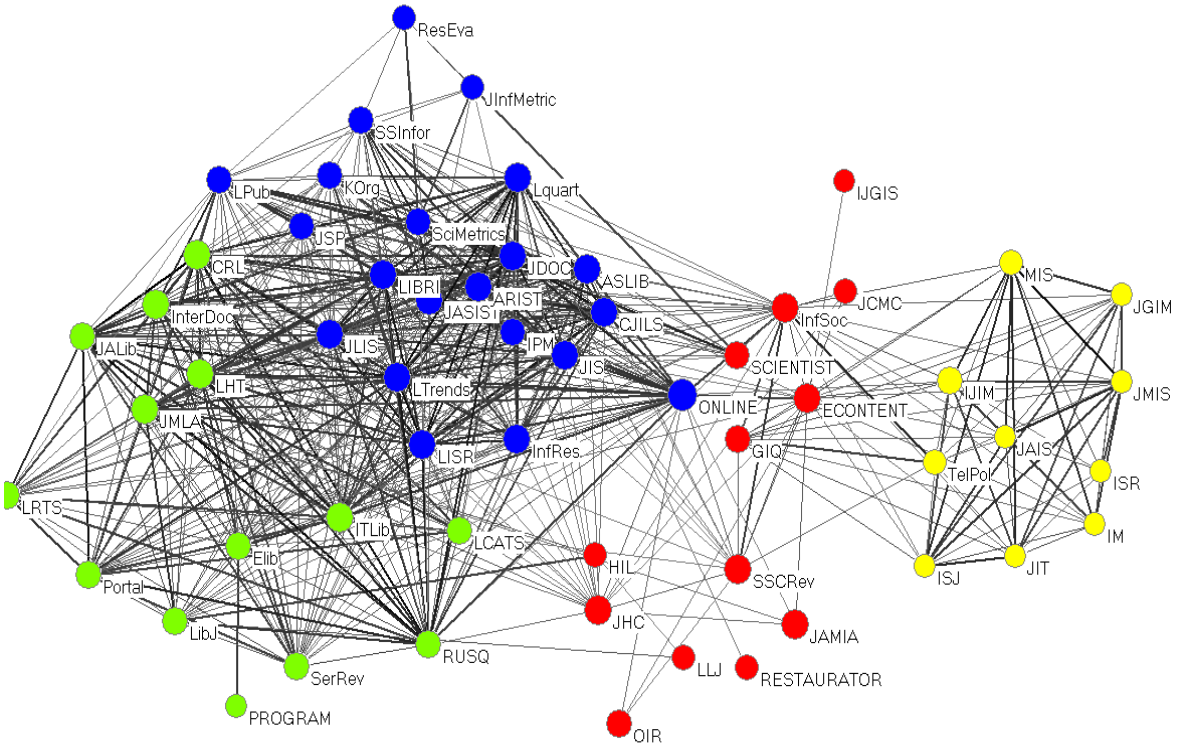

Figure 2 Network View of Journal Clusters by Artifacts

CONCEPTS

Concepts were examined using a topic modeling technique, ACT, based on Latent Dirichlet Allocation (LDA). The ACT model identifies the probability that each journal belongs to a certain topic based on article titles and abstracts. For detailed information about this model, see Tang et al. (2008). Using this method, the top 10 keywords for each topic were identified (see Table 2).

Table 2 Top 10 Words for the Five Topics based on the ACT Model

Topic	LS	Library Practice	Communication & Health	MIS	IS
Top 10 words	Information	Library	Health	Information	Information
	Library	Academic	Information	System	Analysis
	Digital	Librarian	Clinical	Technology	Retrieval
	Web	Digital	Medical	Online	Scientific
	Public	University	Cancer	Internet	Citation
	Social	Study	Analysis	Knowledge	Impact
	Management	Web	Electronic	Management	Search
	Search	Electronic	Communication	Analysis	Evaluation
	Content	Reference	Spatial	Mobile	Knowledge
	Online	Access	Public	Model	Bibliometric

Five dominant topics can be inferred from Table 2: 1) library science and systems (LS), 2) library practice, 3) communication and health, 4) management information systems and 5) information science (IS). The cosine similarity of each journal pair was obtained from the probability distributions of each journal over topics, and the journals were then clustered into five groups. Figure 3 displays the journal proximity network (cosine similarity based on topic distribution), where each node is colored according to the clustering result. The MIS group (yellow) is still a separate cluster, though it is more closely related to the Specialized and communication/health journals (red) than in the previous networks. IS journals (blue) and LS journals (pink) are in slightly different positions than in the producer and artifact networks. An examination of the scope notes of the LS journals reveals that most publish on the topic of digital issues in librarianship. The green nodes are journals mainly related to library service and practice, closely connected to Specialized and LS journals.

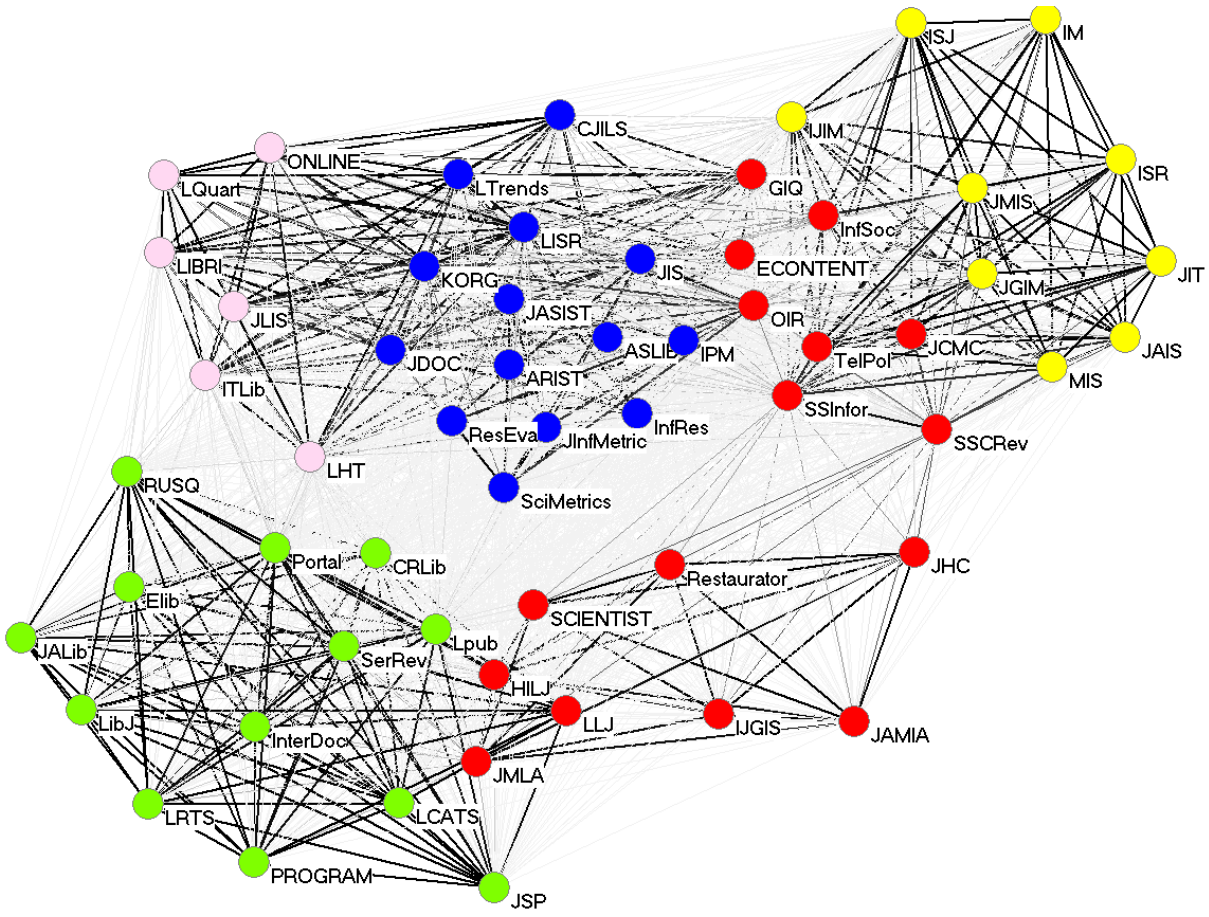

Figure 3 Network View of Journal Clusters by Concepts

GATEKEEPERS

Some 1,561 individual occupied the 1,785 available editorial seats for the 58 IS&LS journals. On average, each journal had 31 editorial board members (s.d.=21.33) and each scholar served on an average of 1.1

editorial boards. A third of the journals had 20-30 scholars serving on their boards. Only one, *Information & Management (IM)*, had more than 100 editorial board members, while one, *Journal of the Association for Information Systems (JAIS),* had more than 80 editorial board members. Thirty-eight scholars serve on three or more editorial boards with the maximum number of editorial board positions held by an individual scholar being 6. Ninety percent of the editorial board members serve on a single editorial board. The interlocking network is thus generated from the 10% of editorial board members serving on more than one journal—demonstrating a high concentration for this variable that differentiates it from other variables.

Ten journals (pink in Figure 4) do not share editorial board members with other journals, that is, they are isolates in the network[5]. The rest belong to four clusters, where the nine MIS journals in yellow constitute a single cluster. *JASIST* and *ARIST* remain at the center. A new cluster emerged in the gatekeeper analysis, consisting of *The Information Society (InfSoc), Journal of Health Communication (JHC), Journal of Computer-Mediated Communication (JCMC), Social Science Computer Review (SSCR)* and *TelPol*, which are all communication-related research journals.

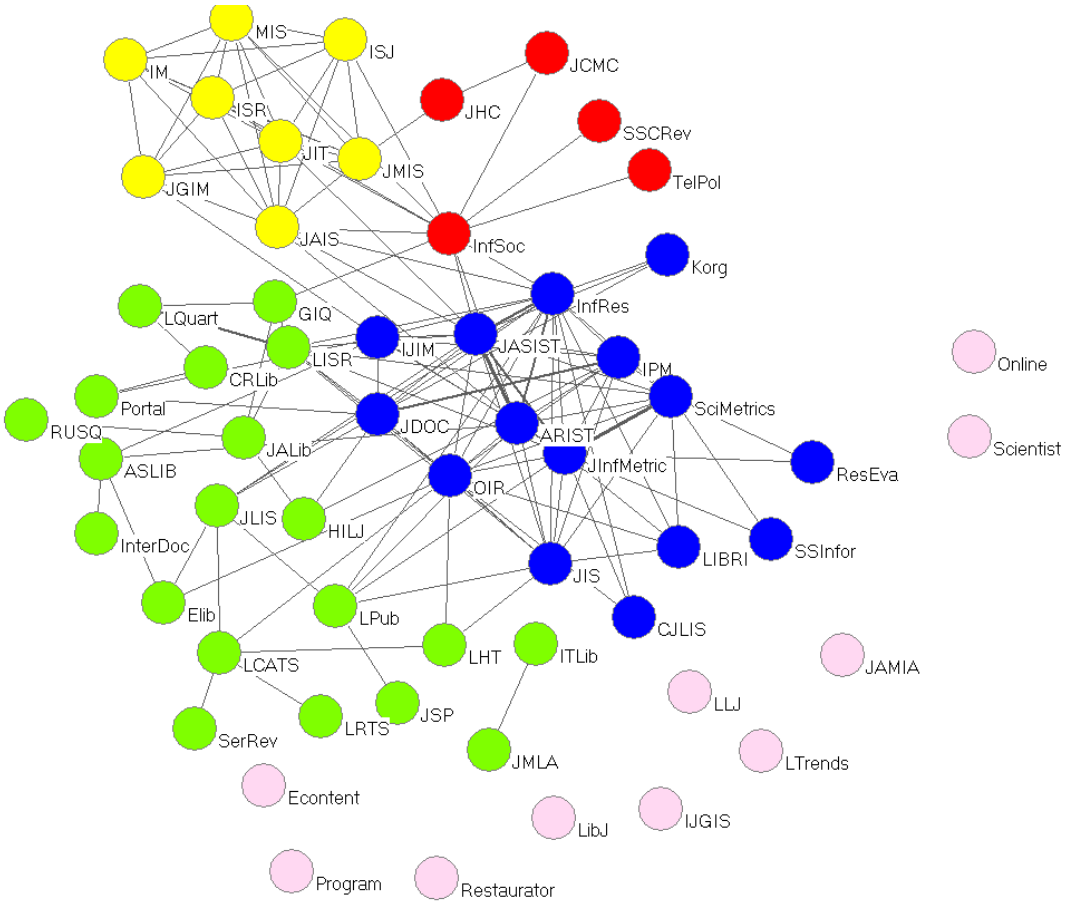

Figure 4 Network View of Journal Clusters by Gatekeepers

[5] Detailed analysis of the reason why these 10 journals do not share editorial board members is not included due to space. Please refer to Ni and Ding (2010).

QUADRATIC ASSIGNMENT PROCEDURE

To examine the relationship among the four facets, a Quadratic Assignment Procedure correlation was performed (Table 3). The QAP correlation test is commonly used in social network analysis as a way of testing the correlation between different networks: multiple networks with the same set of nodes but different relationships. While some typical correlation tests (e.g., Pearson's r) are used to determine relationships between facets, QAP can be used to test the correlation between networks.

Table 3 QAP Comparison of Four Facets

	Co-citation (Artifact)	VAC (Producer)	Editor (Gatekeeper)	Topic (Concept)
Co-citation (Artifact)	1			
VAC (Producer)	0.444*	1		
Editor (Gatekeeper)	0.708**	0.607*	1	
Topic (Concept)	0.421	0.410	0.251	1

**significant at .01 level
* significant at .05 level

The strongest correlation was between gatekeeper (interlocking editorships) and artifacts (co-citation). Also significant were the relationships between gatekeepers and producers (VAC). Significant, albeit to a lesser degree was the relationship between producers and artifacts. Topic organization was not significantly related to artifacts, producers, or gatekeepers.

DISCUSSION AND CONCLUSION

The approaches described here draw upon four facets of the scholarly communication process: a scholar chooses words to describe the content of a work, citations are selected to credit the prior work upon which the article draws, authors submit their work to a preferred journal, and editorial board members/gatekeepers select articles for publication that reflect the scope and standards of the journal they represent (Borgman 1989; Braun 2005). These processes are captured and operationalized using topic analysis, citation analysis, VAC, and editorial board coupling analysis. Similarities and differences in journal clustering by each approach are evident in the results.

First, there are sets of journals that consistently group together across all four facets. Nine MIS journals cluster together and are relatively isolated in each of the networks. The distinct clustering of MIS journals suggests that their inclusion in the IS&LS JCR category should be reviewed. A set of core IS journals also appears consistently together (i.e., *ARIST, CJILS, IPM, JASIST, JDOC, JInfMetrics, JIS, Korg, ResEva,* and *SciMetrics*). The same holds true for a smaller group of LS journals (i.e., *CRLib, ELib, Interdoc, JALib, LCATS, LRTS, and PORTAL*). A smaller group of communication-related journals also appear together across all four facets (*JCMC, JHC, SSCRev,* and *InfSoc*). *LISR* does not cluster consistently with any group of journals, but functions as a bridging journal.

The QAP results indicate that the gatekeeper and co-citation data generate significantly similar journal proximity networks. One possible explanation is that those who serve on the largest number of editorial

boards are also those who dominate the citation landscape. Those who publish the most are more likely to be cited and thus more likely to be chosen as an editorial board member. A comparison of the names of the scholars serving on the most editorial boards by Ni and Ding (2010) and a list of the most cited authors within IS&LS shows some overlap: Thelwall, Cronin, and Davenport appear in both lists. These lists also demonstrate overlap with those who publish the most (Ni et al., to appear), reinforcing the significant relationship between gatekeepers and producers and between producers and artifacts. However, save for a few, the majority of highly-ranked scholars varies by facet.

There were no significant correlations between concepts and any of the other three facets. The weak correlation with topical similarity might provide further evidence for what has long been noted: "Citation is colored by a multitude of factors, not all of which have to do with the accepted conventions of scholarly publishing" (Cronin 1984, p. 31). In other words, citation is not necessarily nor exclusively topic driven.

Gatekeepers and topics have the weakest correlation. Since we assume that interlocking editorial board membership indicates scope similarity (e.g., Baccini & Barabesi 2011), and that topic (based on title words of articles) would also reflect conceptual similarity, these two facets might be expected to provide similar proximity structure for the journals. This is not the case. One possible explanation is that the appointment of editorial board members does not depend solely on their areas of expertise, but also on political and personal factors (Barzilai-Nahon 2009). Our research demonstrates that any single-facet measure of the scholarly landscape will produce partial results and that the utilization of multiple facets is necessary for a full understanding of the scholarly communication system (see, for example, Yan and Ding's [to appear] comparative study on six different scholarly networks). Although correlated, each facet produces a distinct view on the relationship between facets. In a metrics-conscious, indicator-driven environment, those conducting research assessments would do well to note these biases and the potential for error. Future research should focus on ways in which we can more appropriately combine facets for holistic descriptions of the networks of scholarly communication.

ACKNOWLEDGMENT

We are grateful to two anonymous referees for their comments and suggestions.

CITED REFERENCES

Baccini, A., & Barabesi, L. (2011). Seats at the table: The network of the editorial boards in information and library science. *Journal of Informetrics, 5*(3), 382–391.

Bakker, P., & Rigter, H. (1985). Editors of medical journals: Who and from where. *Scientometrics, 7*(1), 11-22.

Barzilai-Nahon, K. (2009). Gatekeeping: A critical review. *Annual Review of Information Science and Technology, 43*(1), 1-79.

Beaver, D., & Rosen, R. (1978). Studies in scientific collaboration Part I. The professional origins of scientific co-authorship. *Scientometrics, 1*, 65-84.

Beaver, D., & Rosen, R. (1979a). Studies in scientific collaboration Part II. Scientific co-authorship, research productivity and visibility in the French scientific elite, 1799–1830. *Scientometrics, 1*, 133-149.

Beaver, D., & Rosen, R. (1979b). Studies in scientific collaboration Part III. Professionalization and the natural history of modern scientific co-authorship. *Scientometrics, 1*, 231-245.

Bedeian, A. G., Van Fleet, D. D., & Hyman, H. H. (2009). Scientific achievement and editorial board membership. *Organizational Research Methods, 12*(2), 211.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research, 3*, 993-1022.

Borgman, C. L. (1989). Bibliometrics and scholarly communication. *Communication Research, 16*(5), 583.

Borgman, C. L., & Furner, J. (2002). Scholarly Communication and bibliometrics. *Annual Review of Information Science and Technology, 36*(1), 1-53.

Braun, T. (2005). Keeping the gates of science journals. *Handbook of Quantitative Science and Technology Research*, 95-114.

Braun, T., & Bujdosó, E. (1983). Gatekeeping patterns in the publication of analytical chemistry research. *Talanta, 30*(3), 161-167.

Braun, T., & Dióspatonyi, I. (2006). Gatekeeping in the international journal literature of chemistry. *Information Processing & Management, 42*(6), 1652-1656.

Budd, J. M. (2000). Scholarly productivity of U.S. LIS Faculty : An update. *The Library Quarterly, 70*, 230-245.

Callon, M., Courtial, J. P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Social Science Information, 22*(2), 191.

Cole, F. J., & Eales, N. B. (1917). The history of comparative anatomy: Part 1-a Statistical analysis of the literature. *Science Progress, 11*, 587-596.

Cronin, B. (2009). Editorial. A seat at the table. *Journal of the American Society for Information Science and Technology, 60*(12), 2387-2387.

Ding, Y. (2011). Topic-based PageRank on author cocitation networks. *Journal of the American Society for Information Science and Technology, 62*(3), 449-466.

Egghe, L., & Rousseau, R. (1990). *Introduction to informetrics*. Amsterdam: Elsevier Science Publishers.

Fox, M. F. (2008). Collaboration between science and social science: Issues, challenges, and opportunities. *Research in Social Problems and Public Policy, 16*, 17-30, doi:10.1016/S0196-1152(08)16001-X.

Garfield, E. (1955). Citation indexes for science. A new dimension in documentation through association of ideas. *Science, 122*, 108-111.

Glogoff, S. (1988). Reviewing the gatekeepers: A survey of referees of library journals. *Journal of the American Society for Information Science, 39*(6), 400-407.

Gross, P. L. K., & Gross, E. M. (1927). College libraries and chemical education. *Science, 66*(1713), 386.

He, T. (2009). International scientific collaboration of China with the G7 countries. *Scientometrics, 80*(3), 571-582.

Jarneving, B. (2007). Complete graphs and bibliographic coupling: A test of the applicability of bibliographic coupling for the identification of cognitive cores on the field level. *Journal of Informetrics, 1*(4), 338-356.

Kessler, M. (1963a). Bibliographic coupling between scientific papers. *American documentation, 14*(1), 10-25.

Kessler, M. (1963b). Bibliographic coupling extended in time: Ten case histories* 1. *Information storage and retrieval, 1*(4), 169-187.

Kessler, M. (1965). Comparison of the results of bibliographic coupling and analytic subject indexing. *American documentation, 16*(3), 223-233.

Lawler, E. L. (1963). The quadratic assignment problem. *Management Science*, 586-599.

Leydesdorff, L. (1997). Why words and co-words cannot map the development of the sciences. *Journal of the American Society for Information Science, 48*(5), 418-427.

Liu, X., Bollen, J., Nelson, M. L., & Van de Sompel, H. (2005). Co-authorship networks in the digital library research community. *Information Processing & Management, 41*(6), 1462-1480.

Luukkonen, T., Persson, O., & Sivertsen, G. (1992). Understanding patterns of international scientific collaboration. *Science, Technology, & Human Values, 17*(1), 101.

McCain, K. W. (1991). Mapping economics through the journal literature: An experiment in journal cocitation analysis. *Journal of the American Society for Information Science, 42*(4), 290-296.

Milojević, S., Sugimoto, C. R., Yan, E., & Ding, Y. (2011). The cognitive structure of Library and Information Science: Analysis of article title words. *Journal of the American Society for Information Science and Technology, 62*(10), 1933-1953.

Ni, C., & Ding, Y. (2010). Journal clustering through interlocking editorship information. *Proceedings of the American Society for Information Science and Technology, 47*(1), 1-10.

Ni, C., & Sugimoto, C. R. (2011). Four-facets of scholarly communities: artifact, producer, concpet and gatekeeper. In the Proceedings of Annual Meeting for the American Society for Information Science & Technology, New Orleans.

Ni, C., Sugimoto, C. R., & Jiang, J. (to appear). Venue-Author-Coupling: A Novel Measure of Identifying Disciplines through Author Communities. *Journal of the American Society for Information Science and Technology*.

Paisley, W. (1990). The future of bibliometrics. In C. L. Borgman (Ed.), *Scholarly communication and bibliometrics* (pp. 281.299): Sage.

Rip, A., & Courtial, J. P. (1984). Co-word maps of biotechnology: An example of cognitive scientometrics. *Scientometrics, 6*(6), 381-400.

Sengupta, I. (1992). Bibliometrics, informetrics, scientometrics and librametrics: an overview. *Libri, 42*(2), 75-98.

Small, H. (1973). Co-Citation in Scientific Literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science, 24*, 265-269.

Small, H., & Koenig, M. (1977). Journal clustering using a bibliographic coupling method. *Information processing & management, 13*(5), 277-288.

Sugimoto, C. R., Li, D., Russell, T. G., Finlay, S. C., & Ding, Y. (2011). The shifting sands of disciplinary development: Analyzing North American Library and Information Science dissertations using latent Dirichlet allocation. *Journal of the American Society for Information Science and Technology*.

Sugimoto, C. R., Pratt, J. A., & Hauser, K. (2008). Using field cocitation analysis to assess reciprocal and shared impact of LIS/MIS fields. *Journal of the American Society for Information Science and Technology, 59*(9), 1441-1453.

Tang, J., Jin, R., & Zhang, J. A topic modeling approach and its integration into the random walk framework for academic search. In *ICDM 2008* (pp. 1055-1060): IEEE

Teichert, T., Heyer, G., Schöntag, K., & Mairif, P. (2011). Co-word analysis for assessing consumer associations: A case study in market research. *Affective Computing and Sentiment Analysis*, 115-124.

Thijs, B., & Glänzel, W. (2010). A structural analysis of collaboration between European research institutes. *Research Evaluation, 19*(1), 55-65.

Van Rijnsoever, F. J., Hessels, L. K., & Vandeberg, R. L. J. (2008). A resource-based view on the interactions of university researchers. *Research Policy, 37*(8), 1255-1266.

Wagner, C. S. (2005). Six case studies of international collaboration in science. *Scientometrics, 62*(1), 3-26.

White, H. D., & Griffith, B. C. (1981). Author cocitation: A literature measure of intellectual structure. *Journal of the American Society for Information Science, 32*, 163-171.

Yan, E., & Ding, Y. (to appear). Scholarly network similarities: How bibliographic coupling networks, citation networks, co-citation networks, topical networks, coauthorship networks, and co-word networks relate to each other. *Journal of the American Society for Information Science & Technology*.

Yan, E., Ding, Y., Milojević, S., & Sugimoto, C. R. (2012). Topics in dynamic research communities: An exploratory study for the field of information retrieval. *Journal of Informetrics, 6*(1), 140-153.

Zhao, D., & Strotmann, A. (2008a). Author bibliographic coupling: Another approach to citation based author knowledge network analysis. *Proceedings of the American Society for Information Science and Technology, 45*(1), 1-10.

Zhao, D., & Strotmann, A. (2008b). Evolution of research activities and intellectual influences in information science 1996–2005: Introducing author bibliographic coupling analysis. *Journal of the American Society for Information Science and Technology, 59*(13), 2070-2086.

Zsindely, S., Schubert, A., & Braun, T. (1982). Editorial gatekeeping patterns in international science journals. A new science indicator. *Scientometrics, 4*(1), 57-68.