

Understanding the Dynamic Scholarly Research Needs and Behavior as Applied to Social Reference Management

Hamed Alhoori and Richard Furuta

Center for the Study of Digital Libraries and
Department of Computer Science and Engineering
Texas A&M University, USA
{alhoori, furuta}@tamu.edu

Abstract. We conducted a study with an objective to learn more about the dynamic information needs, information-seeking behavior, information use and other scholarly activities of researchers. Our focus was on the collaborative and social usage and on the social reference managers. We compared the current practices and strategies of scholars and researchers from multidisciplinary research areas. Our findings provide valuable insights and augment the understanding of how the social web is having a significant effect on the current researchers' activities and digital libraries.

Keywords: Scholarly communication, Research community, Digital Libraries, Information Seeking, Social web, Social reference management, Scholarly references, Social Bookmarking, Knowledge management, Literature review.

1 Introduction

Living and working in a dynamic knowledge society requires deep awareness and access to the best contents and real-time research results to assist and improve strategic decision-making. Understanding the dynamic scholarly activities and their related pathways taken by researchers plays a significant role in developing library collections and services. We have studied several scholarly activities such as: searching, organizing and retrieving articles or notes, collaboration among researchers, usage of social reference managers (SRMs) and its effect on the scholarly activities.

However, while the success of digital libraries increases the materials available to the scholar it also increases the complexity of the scholarly research environment. In this, locally-generated digital libraries serve as a reflection of the frequently-noted global explosion of information—more than 25,000 peer-reviewed research journals exist worldwide, across all disciplines and languages, publishing about 2.5 million articles per year [1]. Filtering and discovering the best results in a short time can be like finding a needle in a haystack.

Traditional libraries provide services to their users while prohibiting them from contributing, and most early digital libraries adopted this viewpoint as well. This results in a considerable loss of external knowledge. However, the current state of the

art is moving toward two ways of interaction, where users both can benefit from the available knowledge and also can contribute to it. Digital library contents moved from being accessed by isolated databases, to a more social and collaborative environments. Instead of being limited by only storing content, personal copies and notes in a personal computer or server, researchers are moving to share and annotate links to their favorite research content on the cloud. Social bookmarking sites [2] have gained visibility in the past few years with these sites reaching thousands of users and containing millions of bookmarks and tags. Beyond increases in effectiveness in finding resources, social bookmarking systems help users to become aware of more relevant information that is available.

Editors and reviewers openly lament the inadequacy of literature reviews in manuscripts submitted for journal publication [3]. Conducting comprehensive research in this era requires having the tools that support the researcher to know the related papers and discover new content. Social bookmarking for researchers [4] or the so-called SRM communities (e.g., citeulike [5] and Mendeley [6]) are playing a significant role in the conduct of research. Retrieving the best results by searching and browsing is no longer the best way to find relevant information.

Algorithms are being designed and developed to suit the scholarly and social community needs such as filtering and discovering items of interest [7], connecting with like-minded researchers and getting recommendations based on their digital libraries content and related work [8]. Other efforts have been directed to reduce the web spam and redundancy [9] that started targeting more specific communities, such as the scholarly world, and introduced a variety of features to fight spam in social bookmarking systems. We have investigated the precision outcomes of a hybrid bibliography system created by an online digital community to support the creation of scholarly bibliographies [10]. Our experimental results indicate that using online reputation based social collaboration improve the quantity and usage of scholarly bibliography and improve the quality and creditability of social citations sites.

This paper is structured as follows. We discuss the related work in Section 2. Section 3 explains the methodology we used. We present and discuss our findings and results in section 4. In section 5 we conclude and highlight some of the future work.

2 Related Work

Abundant studies have been conducted in various disciplines such as geoscience [11], chemistry [12], agricultural and biological sciences [13], medicine and health sciences [14], public health [15], veterinary medicine [16], law [17] humanities [18] to better understand the dynamic information needs, information-seeking behavior, information use and other scholarly activities of researchers, scientists [19], engineers [20], academic scholars [21], undergraduate students [22], graduate students [23]. Several methods were used to examine the scholarly activities using quantitative studies (e.g. surveys [24]), qualitative studies (e.g. interviews) [25], ethnographic observational studies [26] or combination. In [27] they used transactional log studies. Brown used a combination of the e-mail survey and content analysis methods [23]. Some studies used citation analysis [28].

Brown [29] investigated the differences and similarities in information-seeking behavior of academic scientists in four disciplines: astronomy, chemistry-biochemistry, mathematics, and physics. All scientists responding scanned the latest issues of journals to keep abreast of current developments in their fields. The mathematicians surveyed indicated an additional reliance on monographs, preprints, and attendance at conferences and personal communication to support their research activities. Hallmark [30] described the methods of access and retrieval of recent journal articles cited by geoscientists and chemists who work in academia, government, and industry. The study found that the majority of scientists had developed effective new patterns of searching for useful references.

In [31] a multi-disciplinary study explored graduate students' information behavior related to their process of inquiry and scholarly activities. They found that their skills and decisions are influenced directly by professors, other students, librarians and Internet usage. Other results were that the lack of sophistication in finding and using resources and course requirements affect students' information behavior and that findings vary across disciplines and between programs. In addition, some graduate students mentioned influences such as difficulty locating information or the need for convenience and speed.

A number of studies showed that researchers are not aware of or familiar with some of the resources, services and electronic search tools available for them through the library and generally do not consult librarians regarding their information needs [13][32]. Part of the difficulties encountered by researchers in using resources appear to stem mainly from a lack of training [33]. The findings in [32] indicate that guidance in the use of library resources and services is necessary to help students meet some of their information requirements. The study found that journals, library books, and textbooks are the most popular sources of information for course work and research, and that students need to be taught how to use available library resources and services.

Hoffmann, et al. [34], found that graduate students wanted to learn about strategies for finding information, bibliographic management tools, and tools for keeping current with scholarly literature. Students preferred online instruction, although in-person workshops were also found to be valuable. Workshops have been held to support researchers' activities such as using particular tools [35] (e.g., bibliographic management software). The authors of [36] created literature review workshops to serve graduate students from a wide range of subject disciplines at a point of shared need. They identified some of the gaps graduate students have in their knowledge about library services. Nicholas, et al. [37] compared the information seeking behavior of the users' logs of four universities using the OhioLINK journal system and found large differences especially between the research and teaching universities.

Niu and Hemminger, et al. [38], surveyed 2,063 academic researchers in natural science, engineering, and medical science from five research universities in the United States to understand different aspects of researchers' information-seeking behavior. Descriptive statistics were reported by institutions to compare differences among universities. Findings reflected the dominant utilization of electronic methods

for searching and accessing scholarly content. Differences in information-seeking behavior amongst universities were not as clear as those amongst disciplines and demographics. A notable trend is that novel forms of scholarly communication such as collaborative information sharing technology are evolving gradually. They expect that this may be the beginning of a more significant transformative change, particularly in sharing information within laboratories or groups or amongst multisite collaborations. Many professors have begun utilizing blogs, wikis and multimedia to communicate with their colleagues or students. Collaborative search systems, academic social bookmarking systems, open shared rankings and reviews, open access journals, and online sharing bibliographic databases and annotations were all examples of new scholarly communication information technologies. The adoption of these was consistent among the respondents across the five universities.

Most studies were limited on a single campus and did not consider the dynamic changes in scholars information needs and behavior, opportunities and challenges of the social web, or were before its emergence; there were no, or limited, ways of sharing, collaboration, connecting researchers, discovering and recommending content. We have investigated how changes in the technologies available to research communities addressing the use of social media can be used to the benefit of researchers, supporting their overall research progress and outcome. Our research questions included:

- How do researchers search, select, and manage their information sources?
- What difficulties researchers are facing during literature review process?
- How SRM influenced the literature review process?
- What are the current scholarly research needs?

3 Methodology

Our study used two methods of data collection: a qualitative research method using interviews and a quantitative method using an online survey. The same set of questions was used as basis for both methods. Before the methods were carried out, seven researchers reviewed the questions dataset before in order to assess its effectiveness and completion time required. Minimal modifications were made based on this feedback. Participation in both studies was confidential and voluntary. Participants were able to withdraw at any time.

We compared the similarities and differences of researchers considering their scholarly activities. In our interviews, a set of eight randomly selected faculty members from different disciplines on campus (see Table1) were invited to participate in personal interviews. Interviews lasted 45-60 minutes. Most of the faculty interviewed, supervised a research group with active researchers. The interviews started by discussing the current practices in the research group using open-ended questions. Then we moved to cover the unanswered questions from our set of questions. Interview sessions were manually transcribed. Transcriptions of all the discussions were manually coded.

Table 1. Scholars IDs and rank

ID	Research Area	Rank
1	Statistics	Associate Professor
2	Petroleum engineering	Assistant Professor
3	Biology	Assistant Professor
4	Management	Assistant Professor
5	Chemical Engineering	Professor
6	Microbiology	Professor
7	Education	Associate Professor
8	Computer Science	Assistant Professor

The survey was sent to different university departments and social reference manager groups. In the survey, samples were random, independent and quite large, so we used statistical hypothesis testing techniques to investigate how using the SRMs affects the research process. We used Pearson's chi-square test (X^2) and Analysis of variance (ANOVA).

4 Results and Discussion

4.1 Interview

Scholars differ in their reading habits, but in general they agreed that they skim the paper first by reading its abstract, conclusion or results section, and then decide if they will read the complete paper. While some get lost while moving between different papers and references, some kept notes and focus on high impact papers. They agreed that they stop working on literature review when they have enough information and literature content starts to repeat itself.

Most scholars said that they come across at least few articles that would add value to their completed or published work if they knew it exist. In line with conclusions that previous studies showed researchers are finding some difficulties locating their needs as P1 commented:

“I know the information is there, but I don't know how to reach it in a short period of time”.

Scholars mentioned a number of difficulties during the literature review process. Several scholars complained about the repeated results during the continuous search process as P2 stated:

“I would like to have a way to remove the previously viewed results from my new search results or when checking for new citations. Worse than that, when I get some search results that are already stored in my articles collection or reference manager and I start to view them again since my collection is huge and I can't remember all articles, which is totally useless”.

In saving and organizing articles some scholars were still printing articles, and when asked why they didn't move to use advance ways, they said they have been using it for long time and don't want to jump between several tools as P3 commented:

"I print all the papers I need and organize them using authors' names. Although it may take some time to find what I need, however this way works for me since my graduate school".

Few scholars feel satisfied with organizing their papers and notes using folders and text files, as P5 explained:

"I have been using folders to organize my papers and notes based on projects. I know all my folders and when I need anything, I can go back to the project and to the subfolders".

One scholar was even using general organizing tools as P6 explained:

"I am happy using my old file organizing tool version 1.0."

Several scholars used reference managers and share references among their groups. However, others when asked why they don't use a reference manager most were concerned with the learning curve time and possibility of delaying their work as P6 commented:

"I have used the free reference manger provided by the university library. Although it was good, it needs a license and continuous update which delayed my work especially when I move between several places."

Reference managers are becoming an integral tool during the research progress. When asked why P4 was using a reference manager he explained:

"I have around 12,000 articles and I am daily adding few more. I also share some with other scholars".

Scholars take notes on their printed articles or in reference managers. Others were using some online note-taking sites or emails. Few were even using text files and attaching to them all saved articles, notes or ideas. When asked how she remembers where a paper or saved note is, P1 said:

"I have a strong memory, so I know most of my printed papers and the attached notes".

To keep up to date some researchers do a repeat manual search and were not aware of alerts as P5 stated:

"I repeat some searches from time to time and check if there are any new articles to read. Having a tool that can provide me with my research interests can save me a lot of time".

Most scholars have collaborated with other scholars. A major reason was to expand their knowledge and speed the work progress. They select whom they want to collaborate based on others' reliability and ability to collaborate. Some scholars didn't know how the SRM works, and they didn't want to spend time exploring them as P3 commented:

"I am busy with my work and getting my tenure. I don't want to spend time using SRM and adding friends so that I can get the articles recommendations".

However, after knowing how easy SRM works, some were willing to test them and later sends us a thank you letter. A number of researchers expressed regret about their lack of awareness regarding SRMs. P7 was surprised to hear about the SRMs:

"I never heard about the SRM. Actually, I have searched Facebook applications to share my references online with friends, but end up using emails and Google docs".

SRM users showed some concerns about accuracy of the bibliographic data as P8 explained:

"I usually found some errors, missing bibliographic data or duplicate social bookmarks. So, I usually verify its data with the article published press website".

4.2 Survey

156 researchers participated in the online survey as follow (17 faculty members, 5 postdoctoral, 84 doctoral students, 28 master students, 22 undergraduate students). There were 124 male respondents and 32 female; 64% were between 26 and 34 years old. Participants were from 13 different disciplines. We applied several tests to find any significance in the results ($p=p$ -value). We compared how researchers saving methods influence other scholarly activities. Saving methods were using computer folders/directories, reference managers or SRM. We found that SRM users differ significantly from other users in how they search for articles ($X^2=44.31$, $df=4$, $p < 0.001$). While most researchers used general or specific search engines, 40% of the SRM users search within SRM environment. They explained using SRM to search since it has more relevant and newer results, connecting with like-minded researchers or even accurate bibliographic data. SRM users also use tags more often than other users. We found a significant relationship between using SRM and tags usage ($X^2=19.032$, $df=1$, $p < 0.001$). SRM users were able to find more related articles to their research interests than other users. However, there was no significant relationship between using SRM and finding related topics ($X^2=2.11$, $df=1$, $p < 0.05$).

Publications overloading is still a major challenge for most researchers (78%) even for SRM users. However, there was no significant relationship between publications overloading and saving methods ($X^2 = 0.79$, $df=2$, $p < 0.05$) or between publications overloading and how users organize their articles ($X^2=1.35$, $df=1$, $p < 0.05$). Organizing methods were divided into folders, tags or visual tools. Some SRM users show an interest using visual tools but there were no strong evidence of a relationship.

Researchers who use folders get lost more often when reading and moving between articles. We found a significant relationship between saving methods and getting lost while reading and navigating between articles ($X^2=12.71$, $df=6$, $p < 0.05$). We found another significant relationship between saving methods and taking notes on printed papers ($X^2=5.64$, $df=1$, $p < 0.05$). Researchers that take notes on papers constituted 68% of those that use folders, 50% of those that use reference managers and only 19% of those that use SRM. Furthermore, we found a significant relationship between using SRM and taking notes within the SRM ($X^2 = 17.03$, $df=1$, $p < 0.001$).

We found a significant relationship between saving methods and researchers first way to retrieve articles (search or browse) they read recently ($X^2=9.98$, $df=2$, $p < 0.05$). Those that start with search were only 31% of those that use folders, 50% of those that use reference managers and 63% of those that use SRM. There was a significant relationship between saving methods and whether they collaborate with other users or not ($X^2=6.82$, $df=2$, $p < 0.05$). Those that collaborate were 59% of those that use folders, 80% of those that use reference managers and 81% of those that use SRM.

Most researchers collaborate with others (67%) with different reasons mentioned: share and expand knowledge, make new connections, increase possibility of getting funds, motivation, speedup the work or publish more. Researchers that don't collaborate provide different reasons such as, busy with their research, hard to compile/ synchronize the work or don't know other users with similar interests.

Finally, we found a strong evidence that saving methods have an effect on researchers satisfaction while searching ($F=37.80$, $P < 0.001$), while retrieving articles ($F=4.67$, $P < 0.05$) and organizing articles ($F=4.66$, $P < 0.05$).

5 Conclusion and Future Work

This study investigated the current practices and dynamic scholarly activities. It illustrates the remarkable effect that SRMs have had on the scholarly process. The SRM plays a significant role in finding and organizing scholarly articles, connecting researchers, improving collaboration, providing article recommendations, increasing scholarly awareness and revolutionizing scientific communication. SRMs have the opportunity to meet more researchers' information needs and improve their information-seeking behavior.

Academic libraries need to increase the awareness of research technologies available, especially since SRMs are relatively new tools; most users get familiar with a tool and need to be motivated to change to better technologies later. SRM allows the research community to gain many benefits and could have enormous impact in the future on the overall research process.

A 2006 study [31] found that nearly all graduate students (96%) reported that academic staff (e.g., advisers, professors and committee members) influence their research and information seeking. We would like to investigate if SRM has any significant effect on research groups toward building online collaborative research communities. We intend to investigate more the effects of SRMs on the research process and develop a collaborative research model of dynamic strategies. We plan also to investigate how visual tools can influence SRM usage.

References

1. Harnad, S., Brody, T., Vallieres, F., Carr, L., Hitchcock, S., Gingras, Y., Oppenheim, C., Hajjem, C., Hilf, E.: The Access/Impact Problem and the Green and Gold Roads to Open Access: An Update. *Serials Review* 34, 36–40 (2008)
2. Hammond, T., Hannay, T., Lund, B., Scott, J.: Social Bookmarking Tools (I). *DLib Magazine* 11 (2005)
3. Boote, D.N., Beile, P.: Scholars Before Researchers: On the Centrality of the Dissertation Literature Review in Research Preparation. *Educational Researcher* 34, 3–15 (2005)
4. Farooq, U., Song, Y., Carroll, J.M., Giles, C.L.: Social Bookmarking for Scholarly Digital Libraries. *IEEE Internet Computing* 11, 29–35 (2007)
5. Emany, K., Cameron, R.: CiteULike: A Researcher's Social Bookmarking Service. *Ariadne* 51 (2007)
6. Henning, V., Reichelt, J.: Mendeley - A Last.fm For Research? In: 2008 IEEE Fourth International Conference on eScience, pp. 327–328. IEEE, Los Alamitos (2008)
7. Bogers, T., Van Den Bosch, A.: Recommending scientific articles using citeulike. In: *RecSys 2008: Proceedings of the 2008 ACM Conference on Recommender Systems*, pp. 287–290. ACM, New York (2008)
8. Dicheva, D., Dichev, C.: Finding Resources and Collaborators within Digital Collections. In: *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, vol. 1, pp. 631–638 (2010)
9. Heymann, P., Koutrika, G., Molina, H.G.: Fighting Spam on Social Web Sites: A Survey of Approaches and Future Challenges. *IEEE Internet Computing* 11, 36–45 (2007)
10. Alhoori, H., Alvarez, O., Furuta, R., Muñiz, M., Urbina, E.: Supporting the creation of scholarly bibliographies by communities through online reputation based social collaboration. In: Agosti, M., Borbinha, J., Kapidakis, S., Papatheodorou, C., Tsakonias, G. (eds.) *ECDL 2009. LNCS*, vol. 5714, pp. 180–191. Springer, Heidelberg (2009)
11. Bichteler, J., Ward, D.: Information-seeking behaviour of geoscientists. *Special Libraries* 80, 169–178 (1989)
12. Davis, P.M.: Information-seeking behavior of chemists: A transaction log analysis of referral URLs. *Journal of the American Society for Information Science and Technology* 55, 326–332 (2004)
13. Kuruppu, P.U., Gruber, A.M.: Understanding the Information Needs of Academic Scholars in Agricultural and Biological Sciences. *The Journal of Academic Librarianship* 32, 609–623 (2006)
14. Davies, K.: The information-seeking behaviour of doctors: a review of the evidence. *Health Information and Libraries Journal* 24, 78–94 (2007)
15. Revere, D., Turner, A.M., Madhavan, A., Rambo, N., Bugni, P.F., Kimball, A., Fuller, S.S.: Understanding the information needs of public health practitioners: a literature review to inform design of an interactive digital knowledge management system. *Journal of Biomedical Informatics* 40, 410–421 (2007)
16. Pelzer, N.L., Wiese, W.H., Leysen, J.M.: Library use and information-seeking behavior of veterinary medical students revisited in the electronic environment. *Bulletin of the Medical Library Association* 86, 346–355 (1998)
17. Makri, S.: Investigating the information-seeking behaviour of academic lawyers: From Ellis's model to design. *Information Processing & Management* 44, 613–634 (2008)
18. Barrett, A.: The Information-Seeking Habits of Graduate Student Researchers in the Humanities. *The Journal of Academic Librarianship* 31, 324–331 (2005)
19. Tenopir, C., King, D.W., Boyce, P., Grayson, M., Zhang, Y., Ebuon, M.: Patterns of Journal Use by Scientists through Three Evolutionary Phases. *DLib Magazine* 9 (2003)
20. Hertzum, M., Pejtersen, A.M.: The information-seeking practices of engineers: searching for documents as well as for people. *Information Processing & Management* 36, 761–778 (2000)

21. Hemminger, B.M., Lu, D., Vaughan, K.T.L., Adams, S.J.: Information seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology* 58, 2205–2225 (2007)
22. Warwick, C., Rimmer, J., Blandford, A., Gow, J., Buchanan, G.: Cognitive economy and satisficing in information seeking: A longitudinal study of undergraduate information behavior. *Journal of the American Society for Information Science and Technology* 60, 2402–2415 (2009)
23. Brown, C.: Where Do Molecular Biology Graduate Students Find Information? *Science Technology Libraries* 25, 89–104 (2005)
24. Hirsh, S., Dinkelacker, J.: Seeking information in order to produce information: An empirical study at Hewlett Packard Labs. *Journal of the American Society for Information Science and Technology* 55, 807–817 (2004)
25. Bolduc, A.: Surveying user needs in an international context: A qualitative case study from the ILO, Geneva. *The International Information Library Review* 40, 1–9 (2008)
26. Haglund, L., Olsson, P.: The Impact on University Libraries of Changes in Information Behavior Among Academic Researchers: A Multiple Case Study. *The Journal of Academic Librarianship* 34, 52–59 (2008)
27. Nicholas, D., Huntington, P., Jamali, H.R., Watkinson, A.: The information seeking behaviour of the users of digital scholarly journals. *Information Processing & Management* 42, 1345–1365 (2006)
28. Smith, E.T.: Assessing collection usefulness: An investigation of library ownership of the Resources Graduate Students Use. *College & Research Libraries* 64, 344–355 (2003)
29. Brown, C.M.: Information seeking behavior of scientists in the electronic information age: Astronomers, chemists, mathematicians, and physicists. *Journal of the American Society for Information Science* 50, 929–943 (1999)
30. Hallmark, J.: Access and Retrieval of Recent Journal Articles: A Comparative Study of Chemists and Geoscientists. *Issues in Science and Technology Librarianship* 40 (2004)
31. George, C., Bright, A., Hurlbert, T., Linke, E.C., St. Clair, G., Stein, J.: Scholarly use of information: graduate students' information seeking behaviour. *Information Research* 11, 1–19 (2006)
32. Fidzani, B.T.: Information needs and information-seeking behaviour of graduate students at the University of Botswana. *Library Review* 47, 329–340 (1998)
33. Vibert, N., Rouet, J., Ros, C., Ramond, M., Deshouillieres, B.: The use of online electronic information resources in scientific research: The case of neuroscience. *Library & Information Science Research* 29, 508–532 (2007)
34. Hoffmann, K., Antwi-Nsiah, F., Feng, V., Stanley, M.: Library Research Skills: A Needs Assessment for Graduate Student Workshops. *Issues in Science and Technology Librarianship* 53 (2008)
35. Harrison, M., Summerton, S., Peters, K.: EndNote training for academic staff and students: the experience of the Manchester Metropolitan University Library. *New Review of Academic Librarianship* 11, 31–40 (2005)
36. Rempel, H.G., Davidson, J.: Providing Information Literacy Instruction to Graduate Students through Literature Review Workshops. *Issues in Science and Technology Librarianship* 53 (2008)
37. Nicholas, D., Huntington, P., Jamali, H.R.: Diversity in the Information Seeking Behaviour of the Virtual Scholar: Institutional Comparisons. *The Journal of Academic Librarianship* 33, 629–638 (2007)
38. Niu, X., Hemminger, B.M., Lown, C., Adams, S., Brown, C., Level, A., McLure, M., Powers, A., Tennant, M.R., Cataldo, T.: National study of information seeking behavior of academic researchers in the United States. *Journal of the American Society for Information Science and Technology* 61, 869–890 (2010)