How do Computer Scientists Use Google Scholar?: A Survey of User Interest in Elements on SERPs and Author Profile Pages

Jaewon Kim, Johanne R. Trippas, Mark Sanderson, Zhifeng Bao, and W. Bruce Croft

RMIT University, Melbourne, Australia
{jaewon.kim, johanne.trippas, mark.sanderson, zhifeng.bao, bruce.croft}@rmit.edu.au

Abstract. In this paper, we explore *user interest* in elements on Google Scholar's search engine result pages (SERPs) and author profile pages (APPs) through a survey in order to predict search behavior. We investigate the effects of different query intents (*keyword* and *title*) on SERPs and research area familiarities (*familiar* and *less-familiar*) on SERPs and APPs. Our findings show that user interest is affected by the respondents' research area familiarity, whereas there is no effect due to the different query intents, and that users tend to distribute different levels of interest to elements on SERPs and APPs. We believe that this study provides the basis for understanding search behavior in using Google Scholar.

Keywords: Academic search · Google Scholar · Survey study

1 Introduction

Academic search engines (ASEs) such as Google Scholar[1] and Microsoft Academic[2], which are specialized for searching scholarly literature, are now commonly used to find relevant papers. Among the ASEs, Google Scholar covers about 80-90% of the publications in the world [7] and also provides individual profiles with three types of citation namely, total citation number, h-index, and h10-index. Although the Google Scholar website[3] states that the ranking is decided by authors, publishers, numbers and times of citation, and even considers the full text of each document, there is a phenomenon called the *Google Scholar effect* [11]. The Google Scholar effect refers to authors picking and citing papers on the top results on the search engine result pages (SERPs) by assuming these works' credibility and popularity. The growing importance of this search engine in research raises the issues of how people use Google Scholar and what elements in Google Scholar SERPs cause this behavior.

[1] https://scholar.google.com/
[2] https://academic.microsoft.com/
[3] https://scholar.google.com/intl/en/scholar/about.html

In web search, understanding search behavior and user interest plays an important role to suggest better presentation designs of search results [3, 8–10], and to show how search behavior is affected by user background knowledge and query intents [6, 10, 13]. However, according to several previous works related to ASEs [4, 5, 14, 16], academic search behavior can be different from web search behavior due to the different elements (e.g., citation numbers, authors and publication information), goals (e.g., finding a relevant paper), and user groups (e.g., students and academic faculty/staffs), and there have still been insufficient studies on understanding academic search behavior.

As a preliminary work before exploring the search behavior, we conducted a survey study to get an initial picture regarding the following research questions in using Google Scholar:

- RQ1. Do users have different interests in the *elements* on SERPs and author profile pages (APPs)?

In web search, users may have different interests in the elements in SERPs such as title, URL and snippet, and the different user attention may lead to different search behavior [3, 9]. This question asks if the same is true in academic search. Thus, user interest (attention) is the main measurement in this study.

- RQ2. Is the user interest in the elements affected by the *query intent* and *research area familiarity*?

We identified three frequent actions by hearing from several Google Scholar users about how they use it, which are i) *keyword search*: typing a query to find a relevant result (e.g., "artificial intelligence in games"), ii) *title search*: copying and pasting a particular paper title to explore the information (e.g., "Searching for solutions in games and artificial intelligence"), and iii) *profile search*: exploring an author profile page (APP) to see the publication records, citation information, and co-author list by typing the author's name on search engines or Google Scholar (e.g., "Geoffrey Hinton"). In a previous work [15], those actions were also addressed as main categories of query of one academic search engine.

Considering these three actions, we investigated user interests in elements not only in SERPs, but also in APPs. In addition, we adopted *query intent* as a research variable with the two frequent actions (i.e., keyword and title search) to explore the effect on user interest in elements on SERPs. According to several previous studies (e.g., [1, 3, 8–10]), different search purposes produce different search behavior in web search. Although the concept of query intents in this study is somewhat different from the web search taxonomy that classified the search purpose behind the query, the frequent two actions in using Google Scholar may lead to different patterns of user interest.

User background knowledge also leads to different web search behavior. According to results from Kelly and Cool [6], search efficiency increases and reading time decreases as users have higher topic familiarity. White et al. [13] also suggested that domain experts use different strategies and successfully find more relevant results than non-experts do. Therefore, we studied the effect of *research area familiarity* (i.e., familiar and less-familiar) as another research variable.

2 Literature review

There has been some previous research on search behavior in Google Scholar. A few studies investigated the usability and search result quality between Google Scholar and other library systems. A study from Zhang [16] focused on the usability of Google Scholar by comparing a discovery layer system, i.e., Ex Libris Primo. He prepared three pre-defined tasks and allowed one free-typing topic from users, and explored the rating from the relevant judgments using a 7-point Likert scale. The results suggest that Google Scholar received higher usability and preference ratings, and the prepared search results recorded higher relevancy on the search results. Other research compared the quality of sources between using Google Scholar and a library (federated) search tool [4]. She recruited a range of undergraduate students and asked them to identify four relevant sources (i.e., one book, two articles, and one of their interests), related to a self-selected research topic amongst six pre-defined. Her findings indicate that Google Scholar is better for book finding whereas the federated search tool is more useful for searching articles and additional sources.

Some researchers investigated the effects of different users in using Google Scholar. Herrera [5] conducted an exploratory study, where the research variables contained disciplines and types of users, using data from various sources including a Google Scholar library links profile. She found that Google Scholar is mainly used by people in sciences and social sciences disciplines, and graduated students and academic faculty/staffs are the most frequent users. We considered this result when we recruited the respondents. Wu and Chen [14] explored the graduate students' behavior in perceiving and using Google Scholar. Their findings suggest that graduate students generally prefer the usability of Google Scholar than library databases, though their preference was different according to their fields of study.

Although those works generally indicate that academic search behavior can be different from web search behavior due to different types of contents, search goals and users, we currently have insufficient information to understand how users use academic search engines. Therefore, as a preliminarily work for the investigation of academic search behavior, we conducted a survey to explore user interest along elements on SERPs and APPs with the effects of query intent and research area familiarity.

3 Survey study design

3.1 Respondents

We recruited 30 respondents (25 male) via group emailing-lists and a social network. The respondents were required to have a research experience related to

Computer Science in order to obtain responses from more active ASEs users (i.e., graduated students and academic faculty/staffs rather than undergraduates, and sciences researchers rather than all disciplines, as the main users of Google Scholar) by considering the results from Herrera [5]. In addition, respondents must be over 18 years old, and they were required to use a desktop or a laptop. The reason that we recruited respondents from a particular pool (i.e., computer science researchers) is to provide appropriate questions related to less-familiar research areas by assuming that Google Scholar users rarely look up papers totally unrelated to their research areas and by considering the difficulty of preparing the SERPs and APPs of less-familiar research areas for all disciplines. We describe the questions of less-familiar research area in more detail at the next subsection 3.2

The respondents were aged from 18 to 64 and have various educational experience with bachelor (23%), master (17%) or PhD (60%) degrees. 90% of the respondents are working/studying at universities, and over 70% of the respondents replied that they use ASEs and read a paper more than a few times a week. Two-thirds of the respondents identified that they have used ASEs over five years, and all feel confident in using ASEs.

3.2 Questionnaire

Using Qualtrics[4], the survey questionnaire consisted of 22 questions including 14 of consent, qualifying, demographic, and experience questions as briefly reported in the results in the previous subsection 3.1. The remaining eight questions include a question about the frequent actions in using Google Scholar, one of selecting the least familiar research area and six questions to answer the research questions.

To obtain the answer for RQ2 - effects of query intents and research area familiarities, we prepared two question sets about familiar and less-familiar research areas. Each question set contained three questions of keyword and title search on SERPs and author profile search on APPs. To measure user interest in each question for RQ1, the respondents were required to reply about the levels of their interests in each element on SERPs and APPs using a 7-point Likert scale (1: extremely uninteresting, 7: extremely interesting).

As shown in Figure 1, we classified the elements on SERPs as three *element-groups* according to the similarity of their information as content, publication and additional information. In APPs, we categorized the elements to four element-groups as basic, citation, co-authors, and publication information as can be seen in Figure 2.

Before distributing the survey, we performed a pretest with four volunteers to test whether the survey goes well, the data is collected, and the questions are easy to follow. We could make improvements based on the data collected and opinions from the volunteers.

For the question set related to familiar research areas, the respondents were required to prepare their own keywords, paper title and author name, and were

[4] https://rmit.au1.qualtrics.com/jfe/form/SV_3I8roovJ0D46Vpj

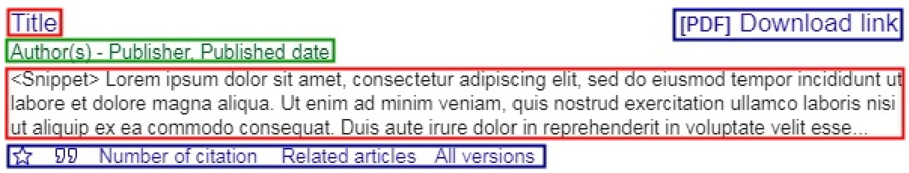

Fig. 1. Sample SERP to guide the respondents to score their interest on each element for the keyword and title search. Each element-group consists of i) content information (red): title and snippet, ii) publication information (green): authors, publisher and published-date, and iii) additional information (blue): number of citation/citing papers, related-articles, all-version and PDF down-loadable.

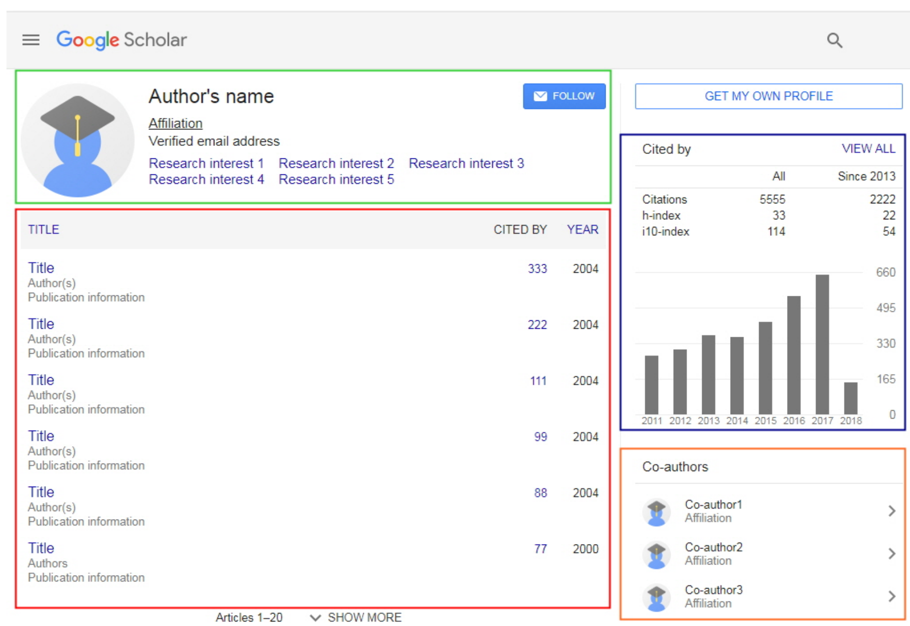

Fig. 2. Sample APP to guide the respondents to score their interest on each element for the profile search. Each element-group consists of i) basic information (green): author's name, affiliation, verified email address, and research interests, ii) publication information (red): paper titles, authors, publication information, cited by, and year, iii) citation information (blue): total citation counts, h/i10-indexes, and citation counts in each year, and iv) co-authors information (orange).

asked to submit them to Google Scholar search embedded in Qualtrics to create SERPs and an APP. For the other question set of less-familiar research area, three pre-extracted/cached SERPs and an APP were automatically presented to the respondents according to their choice of the least familiar research area. The question of selecting the least familiar research area included five categories of computer science research areas (i.e., Big Data and Data Analytics, Information Retrieval and Web Search, Machine Learning and Evolutionary Computing, Intelligent Agents and Multi-Agent Systems, and Networked Systems and Cyber Security). We extracted the keyword SERPs by referring to recent workshop information from top conferences in each research area, prepared the title SERPs by choosing one paper title from the second pages of the keyword SERPs, and obtained APPs by selecting one of the well-known researchers in each research area.

3.3 Design and procedure

In this survey study, we adopted a within-subject design to investigate user interest due to the element-groups (SERPs (3) + APPs(4)) $\times$ query intent (SERPs (2)) $\times$ research area familiarity (2). Thus each respondent took six questions including sub-questions to score their interest regarding each element. To minimize the carry-over effect, we randomized orders of the question sets and the individual questions within the sets, however the orders were counter-balanced across the respondents.

Once the respondents agreed and gave consent, they replied to qualifying, demographic, and experience questions in order. The six questions were then shown to the respondents to ask them to rate the levels of interest in each element.

4 Results and discussion

We obtained 30 data sets from the survey study. We confirmed the power of our design [2] with the significant level $\alpha = 0.05$, that means, the 30 data sets would maintain the power, $1 - \beta \geq 0.95$ for all comparisons in this paper. We focused on analyzing the effects of element-group, query intent (keyword and title) on SERPs and research area familiarity (familiar and less-familiar).

To analyze the score data from the 7-point Likert scale, we adopted a linear mixed model (LMM) [12]. We acknowledge that there may be individual differences in our respondents' pattern of giving scores. To consider this difference, we chose the LMM instead of a linear model (LM) because the observed random effects between the respondents (σ_r^2) were greater than the standard errors (SEs) across the dependent variable - user interest.

4.1 Usage of ASEs and query intents of Google Scholar

We first address the result from one experience question regarding the familiarities with using ASEs. We can observe a significant difference in the familiarity in using ASEs ($\sigma_r^2 = 0.586$, $X^2 = 35.81$, df = 6, p < 0.001). As shown in Figure 3, Google Scholar has the highest familiarity from users (6.53). This supports that the respondents in the survey who have a computer science research background are used to Google Scholar.

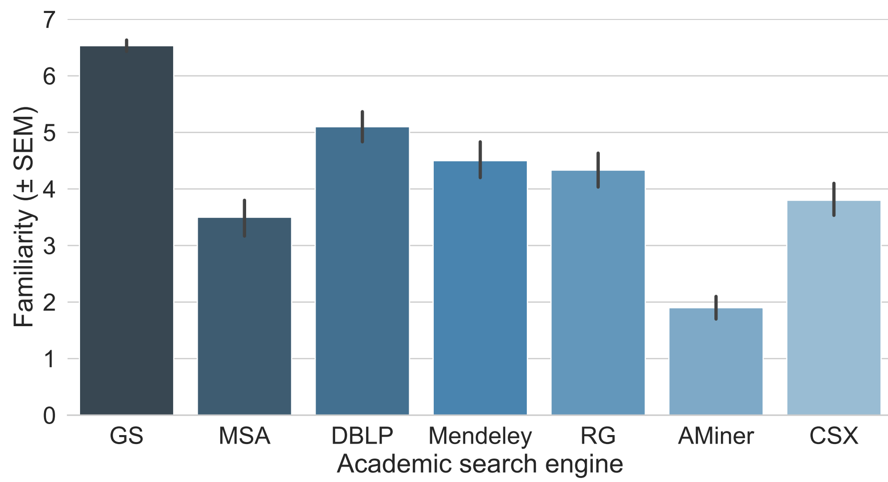

Fig. 3. Users' familiarity in using each ASE. *Note*: GS and MSA denote Google Scholar and Microsoft Academic, and RG, AMiner, and CSX denote Research Gate, Arnetminer and CiteSeerX, respectively.

In addition, we asked the respondents how often they make the actions of keyword, title, and profile search while using Google Scholar, to confirm whether the identified actions are commonly used. Although comparison of frequency of the actions is not necessary, we found a significant effect on the frequency by the actions ($\sigma_r^2 = 0.237$, $X^2 = 8.83$, df = 2, p < 0.001). As shown in Figure 4, the responses for the keyword and title search on SERPs are similar to each other (6.27 and 6.23 for keyword and title search, respectively), whereas the profile search on APPs (5.23) is less used than the others. However, we can confirm that all three actions are frequently made with Google Scholar.

4.2 User interest on SERPs

To test RQ1 —interests in each element-group, we explored the effect on SERPs with two research variables from RQ2 —the query intent and research area familiarity. We found significant effects on user interest according to the element-groups and research area familiarity ($\sigma_r^2 = 0.563$, $X^2 = 38.50$, df = 2, p < 0.001, and $X^2 = 9.84$, df = 1, p < 0.01, respectively), whereas there is no significant difference due to the query intent (p = 0.443).

According to a post-hoc test using standard errors of difference (SEDs), we found the difference between all three element-groups, that is, contents (5.85) $>$ publication (5.19) $>$ additional (4.88) as can be seen in Figure 5. This indicates that the respondents on SERPs are more interested in contents such as title and

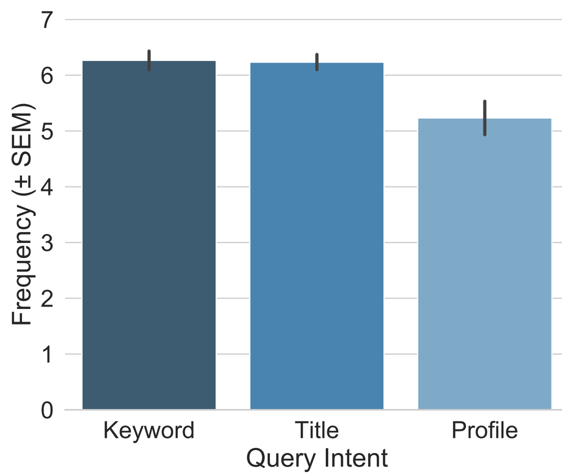

Fig. 4. Frequency of using each query intent.

snippet than the publication information, and they have the least preference for looking at the group-elements of additional information. Relating to the effect of research area familiarity on SERPs, different interests between familiar (5.33) and less-familiar (5.07) research areas were observed, and this suggests that the respondents have more attention in the SERPs extracted related to their familiar research areas.

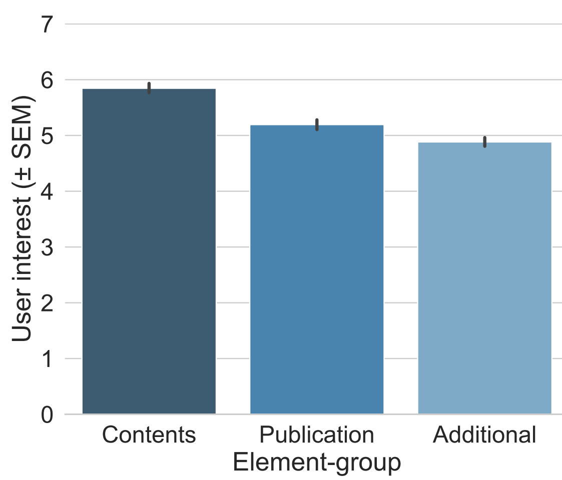

Fig. 5. User interests in each element-group on SERPs.

4.3 User interest on APPs

We then investigated user interest in the element-groups on APPs for RQ1 with the variable of research area familiarity to test RQ2. We can observe significant effects on user interest due to the element-groups, research area familiarity and their interaction ($\sigma_r^2 = 0.328$, $X^2 = 10.95$, df = 3, p < 0.001, $X^2 = 3.92$, df = 1, p < 0.05, and $X^2 = 5.58$, df = 3, p < 0.001, respectively). Our respondents preferred to pay attention to the citation (5.60) and publication (5.50) information rather than co-author (5.00) and basic (5.02) information, and they surprisingly have more interests in APPs with regard to the less-familiar research areas (5.27 and 5.44 for familiar and less-familiar topics, respectively).

To investigate the interaction between two variables, we explored the user interest in element-groups, broken down by research area familiarity as can be seen in Figure 6. Using a post-hoc test by SEDs, we confirmed that the interaction comes from the difference on the basic information (4.66 and 5.38 for the familiar and less-familiar research areas, respectively). That is, the respondents with other element-groups (i.e., publication, citation and co-authors information) on APPs expressed similar interests between familiar and less-familiar researchers' profiles, whereas they tended to be less-interested in basic information while exploring familiar researchers' profile.

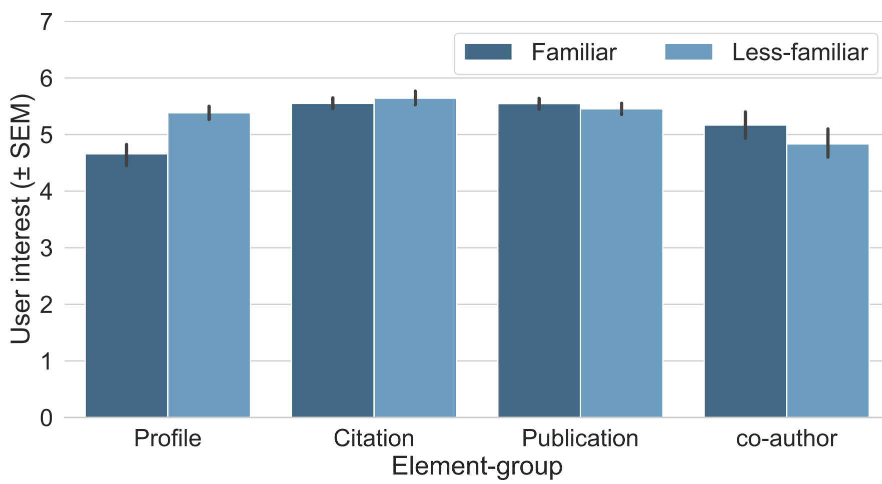

Fig. 6. User interests in each element-group on APPs, broken down by research area familiarity.

5 Conclusions

In this study, we investigated user interest in elements on SERPs and APPs from Google Scholar with considering the effects of query intent and research area familiarity. On SERPs, we found that users are more interested in the content information than other elements, and they tend to have more interests in SERPs with familiar research areas, whereas we could not observe effects of the query intent (i.e., keyword and title). On APPs, the citation and publication information received more attention from the respondents, and the users have less interest in APPs related to familiar research areas. In addition, we confirmed that users recorded lower interests on the basic information when they look at a familiar author's profile.

We acknowledge that this study has several limitations, that is, we recruited people who have a computer science research background, we adopted a particular ASE —Google Scholar, and the results from a survey study can be different from user's actual search behavior. As a preliminary work, this study provides a basic information for search behavior in using Google Scholar. For the future work, we plan to conduct a lab-based eye-tracking user study to explore how user interest moves (i.e., fixation information) by comparing to the survey results and what their decisions are (e.g., clicks and next pages from SERPs and APPs) for better understanding of using Google Scholar.

Acknowledgements This work was partially supported by the Australian Research Council's discovery Project Scheme (DP170102726).

References

1. Broder, A.: A taxonomy of web search. ACM SIGIR Forum 36(2), 3–10 (2002)
2. Chow, S.C., Shao, J., Wang, H., Lokhnygina, Y.: Sample size calculations in clinical research. Chapman and Hall/CRC (2017)
3. Cutrell, E., Guan, Z.: What are you looking for?: an eye-tracking study of information usage in web search. In: Proceedings of the SIGCHI conference on Human factors in computing systems. pp. 407–416. ACM (2007)
4. Georgas, H.: Google vs. the library (part iii): Assessing the quality of sources found by undergraduates. Portal: Libraries and the Academy 15(1), 133–161 (2015)
5. Herrera, G.: Google scholar users and user behaviors: an exploratory study. College & Research Libraries 72(4), 316–330 (2011)
6. Kelly, D., Cool, C.: The effects of topic familiarity on information search behavior. In: Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries. pp. 74–75. ACM (2002)
7. Khabsa, M., Giles, C.L.: The number of scholarly documents on the public web. PloS one 9(5), e93949 (2014)
8. Kim, J., Thomas, P., Sankaranarayana, R., Gedeon, T., Yoon, H.J.: Eye-tracking analysis of user behavior and performance in web search on large and small screens. Journal of the Association for Information Science and Technology 66(3), 526–544 (2015)
9. Kim, J., Thomas, P., Sankaranarayana, R., Gedeon, T., Yoon, H.J.: What snippet size is needed in mobile web search? In: Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval. pp. 97–106. ACM (2017)
10. Lorigo, L., Pan, B., Hembrooke, H., Joachims, T., Granka, L., Gay, G.: The influence of task and gender on search and evaluation behavior using Google. Information Processing & Management 42(4), 1123–1131 (2006)
11. Serenko, A., Dumay, J.: Citation classics published in knowledge management journals. Part ii: studying research trends and discovering the Google Scholar effect. Journal of Knowledge Management 19(6), 1335–1355 (2015)
12. West, B.T., Welch, K.B., Galecki, A.T.: Linear mixed models: A practical guide using statistical software. CRC Press (2014)
13. White, R.W., Dumais, S.T., Teevan, J.: Characterizing the influence of domain expertise on web search behavior. In: Proceedings of the second ACM international conference on web search and data mining. pp. 132–141. ACM (2009)
14. Wu, M.d., Chen, S.C.: Graduate students appreciate Google Scholar, but still find use for libraries. The Electronic Library 32(3), 375–389 (2014)
15. Xiong, C., Power, R., Callan, J.: Explicit semantic ranking for academic search via knowledge graph embedding. In: Proceedings of the 26th international conference on world wide web. pp. 1271–1279. International World Wide Web Conferences Steering Committee (2017)

16. Zhang, T.: User-centered evaluation of a discovery layer system with Google Scholar. In: International Conference of Design, User Experience, and Usability. pp. 313–322. Springer (2013)