Exploratory Search in Digital Libraries: A Preliminary Examination of the Use and Role of Interface Features

Lori McCay-Peet and Anabel Quan-Haase
University of Western Ontario
London, Ontario, Canada
mccay@dal.ca and aquan@uwo.ca

Dagmar Kern
GESIS – Leibniz Institute for the Social Sciences
Cologne, Germany
Dagmar.Kern@gesis.org

ABSTRACT
What interface features of academic digital libraries (DLs) do scholars use when engaging in an exploratory search task? We compare social scientists’ and computer scientists’ use of interface features in domain-specific DLs to demonstrate what features may lead to useful and unexpected findings of information. Applying a search user interface framework, we found social scientists were more likely than computer scientists to use control features (e.g., filter options) during exploration. Both sets of scholars found useful information during exploration, but social scientists more frequently reported they also found something unexpected. Findings suggest differences may be due to variances in the interface features used by the two groups. We discuss future research including further examination of the experiences that the use of interface features may elicit.

Keywords
Information seeking, scholars, academic discipline, exploratory search task, interface features, serendipity.

INTRODUCTION
Academic digital libraries (DLs) are a key information source for scholars (Borgman, 2000). But what do scholars want to discover, learn, and experience in these digital spaces? Marchionini (1995) critically noted that “it remains doubtful whether any single generic ‘window’ on the universe of information will be sufficient to support the views required for all purposes and users” (p. 158). Twenty years later we are still grappling with the problem of the generic search user interface (SUI).

SUIs often present a multitude of features from which users can select on an as-needed basis, including summaries, advanced search, and subject categories. This approach can be detrimental, as it leads to highly complex and cluttered SUIs (Wilson & Schraefel, 2008), and further underlines the importance of understanding the use and role of specific interface features during a task. Even though interface features have readily evident functions—e.g., search boxes are for querying, filters are for narrowing results—the design of an interface “is no straightforward task” (Huurdeman & Kamps, 2014). There is little understanding in the literature about the role users perceive features to play and hence we cannot assume the intended function of a feature is the only one the user experiences.

Despite the various approaches (e.g., heuristics) designers have to guide them in the construction of SUIs, “there is little theoretical work that puts forward a solid explanation to how different search engine and digital library interface features are used or what their role or importance is” (Diriye et al., 2013). This poster helps fill this void through an investigation of the use of various features of DLs during an exploratory search task. We compare two groups of scholars in domain-specific DLs: computer scientists and social scientists. Our examination of interface features reveals how these features may be related to finding useful and unexpected information, experiences associated with learning and serendipity in work and scholarship (McCay-Peet et al., 2015).

Two research questions are investigated:

1) What features of domain-specific DLs do scholars use during an exploratory search task?

2) Are there differences in perceived usefulness and unexpectedness of the information found in two domain-specific DLs?

BACKGROUND
Prior research on the information seeking behaviours of computer scientists indicates reasons for seeking information include “keeping up to date, exploring new topics, reviewing literature, collaborating, preparing lectures, and recommending material for students” (Athukorala et al., 2013). Similarly, social scientists’ information behaviours are motivated by the need to conduct literature searches to find information and sources, maintain awareness of relevant research, network with colleagues in their field, and manage information (Meho & Tibbo, 2003). A portion of these information behaviours involve exploratory search, which “blends querying and browsing strategies” (Marchionini, 2006, p. 41) and is generally associated with sensemaking, uncertainty, and difficulty because problems that drive exploratory search are often complex, multifaceted, and ill-defined (Wildemuth & Freund, 2012).
Through a comparison of search logs of two domain-specific DLs—the National Science DL and Opening History—Zavalina and Vassilieva (2014) found there were differences in the ways in which the two research communities searched (e.g., length and type of search queries). However, the examination was limited to the use of input features (e.g., basic keyword, advanced search) rather than the whole SUI, which the current paper examines. Prior research indicates that the design of SUIs, encompassing not just input features, but features such as recommendations, has an impact on information seeking (Diriye et al., 2010). Moreover, different types of digital environments are more adept at facilitating unexpected experiences with information (e.g., social media versus databases) (McCay-Peet et al., 2015), which suggests that features common to different types of digital environments may play an integral role in user behaviour and experience. This poster seeks to understand what features are associated with useful and unexpected experiences with information.

**METHODS**

Participants, 15 German scholars (10 male; 5 female), primarily PhD students (n = 6) and post docs (n = 6), were asked to complete an exploratory search task to gather information for a grant proposal (n = 14) or term paper (n = 1), related to their own research interests. Computer scientists (n = 8) completed the task in the ACM Digital Library (ACM DL)¹, while social scientists (n = 7) completed the task in Sowiport², a social science-oriented DL. Social scientists reported they used Sowiport (m = 2.00; sd = 1.29) as frequently as computer scientists reported using the ACM DL (m = 2.88; sd = 1.13) [F (1, 13) = 1.97, p > .05]; all ratings were completed on a 1 to 5 scale.

Participants were given ten minutes to engage in the exploratory search task, which was recorded using screen recording software. The post-task questionnaire was administered using Limesurvey and participants rated to what extent they encountered useful and unexpected information. For a baseline, participants were also asked, in general, how frequently they stumbled upon information relating to their research or study interests while searching for something else in a DL—a serendipitous experience relating to both the concepts of useful and unexpected (McCay-Peet et al., 2015).

---

¹ http://dl.acm.org/
² http://sowiport.gesis.org/
The ACM and Sowiport DLs contain similar features, though differences exist in the placement of some of these features and means to, for example, filter and sort. Fig. 1 shows the two main types of pages used by participants. The ACM and Sowiport DLs’ (A) pages to the left exemplify the primary search pages, and those on the right (B) show the type of page to which users were directed when they clicked on a link to an item (e.g., journal article). We employed Wilson’s (2011) SUI framework to analyze the data. The SUI framework was developed based on features available in search engines such as Google, but we found it could be adapted for DLs (Table 1). The types of features on the DL pages are identified in Fig. 1 numerically: (1) Input, (2) Control, (3) Informational, and (4) Personalizable features. Wilson (2011) previously noted overlaps among the feature types and this overlap was apparent in our analysis as well. Illustrating the complexity of the analysis, when participants looked at the references in the ‘cited by’ feature of the ACM DL, we coded this as “Citations” under the feature type of “Informational” because it provided information about the item. However, if participants then clicked on one of these citations, we further coded this portion of the screen recording as the use of a “Similar or related item” feature under the feature type “Control” because the list of citations then went beyond an informational role by enabling users to modify their original search by narrowing their search or providing an opportunity to go in a different direction with their search.

NVivo was used to code the recordings using Wilson’s (2011) SUI framework and SPSS was used to analyze the data together with closed-ended survey responses. Open-ended survey responses were imported into Excel and analysed thematically.

**FINDINGS**

**Use of features**

Table 1 outlines the features used by participants. The majority of the features were classified under one of Wilson’s (2011) four SUI feature types, but four features were categorized as External (i.e., email, record, external website, and external search engine) as they were software applications, and resources accessed outside of the DL.

Participants used a minimum of three features during their exploratory search task and a maximum of eight (m = 6.1). All participants used Results Lists and Summaries (n = 15). All but one participant used Search Box (n = 14). With the exception of the Filter Options (n = 8), the remainder of the features were used by less than half of the participants. Social scientists used more features (m = 7.00; sd = 1.16) in Sowiport than computer scientists (m = 5.25; sd = 1.28) in ACM DL [F (1, 13) = 7.62, p < .05].

<table>
<thead>
<tr>
<th>Type of feature</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Input: features that enable users to explicitly indicate what they are looking for.</td>
<td>Advanced Search, Search Box</td>
</tr>
<tr>
<td>2. Control: features that allow users to modify their original search (e.g., narrow, expand).</td>
<td>Filter Options, Google Scholar*, Similar or related items (e.g., links to cited items, list of author’s references), Sort, Subject Categories, Thesaurus</td>
</tr>
<tr>
<td>3. Informational: features that either give the search results or give the user information about the results or an item (e.g., conference paper).</td>
<td>Citations, Full Text, Results List, Summaries (e.g., abstracts)</td>
</tr>
<tr>
<td>4. Personalizable: features that link users with their prior interactions.</td>
<td>Email*, Favourite Options</td>
</tr>
<tr>
<td>5. External: software, applications, or resources accessed outside of DLs.</td>
<td>Email*, Record (i.e., Microsoft Word, notepad), External Website, External Search Engine</td>
</tr>
</tbody>
</table>

*Option to search via Google Scholar within Sowiport was used by 1 participant. Email classified as External or Personalizable, depending on whether or not the DL email application was used. Features used by more than half of participants are bolded.

**Table 1. Features used during task, classified using adapted SUI framework (adapted from Wilson, 2011).**

Social scientists were more apt to use Control features (m = 4.71; sd = 3.03) than the computer scientists (m = 1.25, sd = 1.58); t(8.76) = -2.71, p < .05. There were no other significant feature type differences. Within Control features, the only significant difference was in the use of Filter Options, which were used more frequently by social scientists (m = 2.71; sd = 2.06) than computer scientists (m = 0.50; sd = 1.07); t(8.75) = -2.56, p < .05.

**Finding useful and unexpected information**

A Mann-Whitney Test indicated that there was no difference between social scientists’ (Mdn = 3) and computer scientists’ (Mdn = 3) self-rated usefulness of the information found in the exploratory search, U = 24.50, p = .63, r = -0.13. However, a Mann-Whitney Test indicated that the self-rated unexpectedness of the information found was greater for social scientists (Mdn = 3) than computer scientists (Mdn = 1), U = 5.50, p = 0.009, r = 0.70 (Fig. 2).

When asked what they did (or did not) find unexpected about the useful information they found, responses primarily related to the depth and breadth of the information found. Two computer scientists reported they found the types of papers they expected to find. One noted, “Found fairly recent papers in good and the expected conferences from the usual people who I would expect to work in the
area” (15CS). One participant who had found articles on a specific area of experimental sociology, found the experience unexpected because “I had expected to find less experiments” (2SS).

To test whether the differences illustrated in Fig. 2 were potentially related to discipline rather than DL, a Mann-Whitney test was conducted. There were no differences in how frequently social scientists ($Mdn = 3$) and computer scientists ($Mdn = 3$) experienced stumbling upon information relating to their research or study interests while searching for something else in DLs (serendipity), $U = 80.5, p = .72, r = -0.09$.

![Figure 2. Found something unexpected in the DL.](image)

**Figure 2. Found something unexpected in the DL.**

**DISCUSSION AND CONCLUSIONS**

The small sample size limits our study’s generalizability, but findings raise a number of questions that point to future research. Our study involved two different domain-specific DLs and thus two different research communities making it difficult to discern whether differences in behavior are due to the DLs or domains, however, our study suggests that while both social and computer scientists are just as likely to experience serendipity in general, differences in their use of features within DLs may play a role in when these experiences occur. Most or all participants used a selection of features during the exploratory task (i.e., search box, results list, summaries, filter options), but the number of features used varied. Social scientists, who were more likely to use filter options and more features in general, were also more likely to indicate they had unexpected experiences with information in the DL, namely related to perceptions of the depth and breadth of research relating to their topic. Does the use of control features, and particularly filter options, lead to more surprising, unexpected interactions with information in digital libraries? Does the use of more features in general lead to a significantly better search experience with more unexpected results? Can the provision and use of a more parsimonious set of features in a DL have the same impact?

Research is needed to understand the use and role of features in domain-specific DLs (Zavalina & Vassilieva, 2014). Understanding the experiences users derive from features (e.g., unexpected, delightful) may provide insights into motivations behind feature use and further inform the design of DLs. While we do not suggest stemming the tide of novel feature development, research does need to examine the uses and roles of features to inform their development and integration into DLs. Exploring other ways in which to study features in DLs would enable a nuanced perspective of various features’ uses and roles.

**ACKNOWLEDGMENTS**

This research is funded by a SSHRC Insight grant given to A. Quan-Haase (No. R3603A13).

**REFERENCES**


Zavalina, O., & Vassilieva, E. V. (2014). Understanding the information needs of large-scale digital library users searching. *LRTS*, 58(2), 84–100