

UC Health Sciences Metasearch Exploration:

Part I: Graduate Student Focus Group Findings

Prepared for: UC Health Sciences Metasearch Team
Report Author: Jane Lee
Report Contributor: Felicia Poe
On-Site Facilitators: Rebecca Davis and Jo Anne Boorkman

DRAFT
Last modified 28 March 2006

Introduction

The California Digital Library is working to create metasearch tools and software that campus libraries can use to craft search portals tailored to specific audiences and needs. A primary goal of the metasearch service is to assist users in efficient discovery of information across a range of resources. The UC health sciences librarians are exploring the possibility of creating a metasearch portal for health sciences faculty, residents, and graduate students.

This document describes the findings from two focus groups that were held on March 7-8, 2006 in the Carlson Health Sciences Library at UC Davis for the first phase (looking at graduate students) of the UC Health Sciences Metasearch Exploration.

Focus Group Findings

Research Practices and Resources

- The big three academic resources: PubMed, Web of Science, BIOSIS.
- Full text is king.

Graduate students in the health sciences primarily use three academic resources: PubMed, Web of Science, and BIOSIS. Participants expressed the following sentiments about these resources:

- PubMed is accessible and exhaustive.
- “I’m a diehard PubMed addict.” UC-eLinks is great.
- “PubMed and Web of Science have everything that I need.”
- PubMed is more user friendly than BIOSIS, and it’s easier to upload to EndNote from PubMed.

Participants reported that when beginning research, they try to find a good article to start with and go from there. They use review articles and look at citations and related articles. In doing research, students desire full text above anything else. Given a choice between full text and an abstract, people will choose full text, even if it’s not exactly what they are looking for. Some prefer PDF’s because they preserve formatting, graphs, and tables. Others prefer HTML because it provides hyperlinked references that are easier to access.

Google, et cetera

- Google is an important resource for graduate students, but it is used for specific purposes.
- Google Scholar = The People’s Metasearch
- Wikipedia

Graduate students rely on Google (<http://www.google.com/>) to satisfy many of their information needs, but they recognize its limitations. One participant observed that “[you] don’t know who’s putting stuff up there.” Another noted that using Google was sometimes frustrating, because there is too much information to sift through, particularly when searching for images. Still, many consider Google as the best resource for quick reference questions and information about particular techniques. One participant expressed a desire for the most up-to-date versus the most comprehensive information. He studies obesity, so the non-scholarly world has a lot of information. “Google completes the circle” by giving the “breaking news story.”

Graduate students gave the following reasons for using Google Scholar (<http://scholar.google.com/>), a beta service of Google aimed at academic researchers:

- “Everything comes up.”
- It’s easier than PubMed.
- It’s sometimes easier to find and get PubMed stuff through Google Scholar than within PubMed.

One user reported that she hadn’t put much effort into figuring out how to access the library’s website through a proxy server, because Google is so easy. She said that with the library, it’s not as easy or appealing or relevant. Instead, she uses Google to get stuff “backdoor”.

One participant said that she also goes directly to authors’ sites if she cannot find the full text of an article through her usual means. Sometimes PDF’s are available there. If not, she contacts the author via email to request a copy of the article. This participant reported a reasonable degree of success with this method.

A number of graduate students also reported using Wikipedia (<http://wikipedia.org/>) in addition to Google. Because Wikipedia’s content can be created and edited by anyone, these students used it mainly to look up definitions and information for coursework, not necessarily research. It is interesting to note, however, that participants who used Wikipedia held it in high regard despite recent incidents of fabrication and falsification of entries. Perhaps because of their faith in the power of the “collective”, these students seemed to consider the integrity of information found on Wikipedia to be greater than that of the general Web.

Information Management

- Print
- Save
- Email
- EndNote
- Alerts
- Basket

Participants reported using several methods to manage information. A common practice was to print articles and organize them into folders. Those who preferred to print hard copies of articles gave the following reasons for doing so:

- If a PDF version is available, this user will save a copy to the hard drive, but his PDF’s are scattered across different computers. It’s easier for him to print articles and keep track of them in file folders.
- One participant likes the ability to take stuff with her and highlight and annotate right on the page.
- One participant reported that if she expects to really use an article, she prints it in order to make notes in margins. If she just wants it for reference, she saves it to her hard drive.

Another common practice was to save PDF’s in appropriately named folders on a hard drive. One participant said that a lack of file cabinet space was the main reason he didn’t automatically print articles. Another participant reported that he didn’t even save PDF’s. Instead, he saves PubMed ID numbers or links to abstracts. Another participant maintains a personal website of resources, so she can access her research materials from anywhere. She uses a FileMaker database to keep track of citations.

Participants reported using their email accounts to store and manage information. Because lack of time is an issue, participants try to keep their email free of unnecessary information. One participant stated that when she's busy, even email alerts that she signed up to receive are annoying and just get deleted.

EndNote

Several participants considered EndNote an essential tool for those dealing with large numbers of citations. In part, graduate students' satisfaction with PubMed is a result of EndNote's ability to import data directly from PubMed. Students also mentioned EndNote's "cite while you write" feature as especially helpful. EndNote is such an integral part of these graduate students' research workflow that the successful adoption of any tool or resource depends on its ability to integrate with EndNote.

One participant uses EndNote to organize the PDF's on her hard drive. Another uses it to "maintain an accurate record of what I have read." A few had never heard of the program before. One person owned it, but had not used it because she was waiting to take a library class on using EndNote.

Alerts

Participants report trying to stay informed about new resources via alert services that email notices to them. These emails usually contain the tables of contents of new resources. For example, users may save searches on PubMed and have new results from their searches emailed to them. One participant reported using PubCrawler (<http://pubcrawler.gen.tcd.ie/>),

a free 'alerting' service that scans daily updates to the NCBI Medline (PubMed) and GenBank databases. PubCrawler helps keeping scientists informed of the current contents of Medline and GenBank, by listing new database entries that match their research interests.

Referring to PubCrawler, this participant stated, "You get the latest and greatest right away." Another participant reported using a service that notifies subscribers when an article is cited.

Basket

Participants did not see a great need for saving items to a basket within session. Most either print desired articles as they find them or save copies onto their hard drives. However, many participants did express interest in an across sessions basket concept. A few mentioned the PubMed clipboard, which saves searches across sessions, thereby allowing its users to keep track of what they've searched and seen. One participant believed that a cross-session basket would help eliminate redundancy. He described a scenario of saving the same great article three times because he didn't remember having done that search before. One point to note is that participants seemed to want the basket to keep track of the searches they performed rather than individual saved items. They already have other methods for keeping track of articles.

Discovery Tools

- Recommendations: Help me correct my mistakes.
- Browsing
- Advanced Search

Recommendations for additional resources received mixed reviews from participants. Some viewed them as helpful; others thought they were distracting. One participant was skeptical about "popular" recommendations, i.e. recommendations based on the popularity of an item with other

patrons. Another participant remarked, "I'm really good at getting myself off track on my own." One participant reported that if she doesn't know what she's looking for, then recommendations are more helpful. If she knows what she is looking for, then she is able to recognize what's good on her own.

All agreed that recommendations that help correct mistakes, such as Google's "Did you mean" feature, were extremely useful.

Most participants reported doing keyword or author searches using basic search interfaces. However, one person said that he preferred searching Web of Science through its advanced search interface rather than through EndNote, because the Web of Science interface offered more options.

Participants did not report doing much browsing, other than viewing tables of content. Most of the time, they report searching for known items.

Metasearch

- "That's why I use Google Scholar."
- Graduate students desire one-click access to full text articles, as well as guidance on which databases to use.

Graduate students reported being uncertain about how a metasearch service would work and what it could offer them. Some saw potential benefits. ("That's why I use Google Scholar.") Others did not see a need for metasearch because, in their opinion, each database has a different purpose.

Participants expressed the following concerns and desires:

- PDF must be available.
- Deduplication of results is required.
- The ability to customize the list of databases searched
- The ability to sort and filter search results
 - By database, e.g. show only articles from PubMed and BIOSIS
 - "Show me what it is and I'll sort it myself on the page."
- The ability to upload to EndNote
- Including books may be useful. Sometimes general information is needed, so getting a chapter from textbook, as well as PubMed articles, may be helpful. Metasearch sounds very appealing from a TA (teaching assistant) perspective.
- The desire to know/learn which databases are relevant. Maybe one could use metasearch to see which databases have relevant content. This may help students discover new databases.
- Google Scholar does metasearch, but if UC's metasearch service saved searches and included Harvest and Melvyl, it could be better than Google Scholar.
- The user interface needs to be usable and customizable.
- The ability to search within results
- "Relevance is the important issue."
- "I would use this if it would yield more relevant articles. But it also has the potential to yield even more irrelevant articles as well." (In other words, the results would probably include a greater *number* of irrelevant articles.)

Appendix A: Questions and Objectives

1. Please introduce yourself and describe your area of expertise.

Objective: Give participants an opportunity to speak in order to break the ice.

2. Is there anything unique about your area of expertise that makes research particularly challenging compared to other fields?

Objective: Determine sources of pain, which might help identify a role for MetaLib.

3. After identifying a research question, how do you get started? What are your strategies for finding information?

Objective: Determine research behaviors of users.

4. Do you prefer to browse by topic or do a keyword search?

Objective: Determine users' preferences for information gathering.

5. [Advanced vs. basic search]

Objective: Determine users' search interface preferences.

6. How do you stay current? What resources do you use?

Objective: Determine whether or not users use non-traditional resources.

7. There are many different kinds of sources – such as databases and journals – that one may use for research. How do you decide which ones to use?

Objective: Determine research behavior of users.

8. What kinds of difficulties do you run into in the course of doing research?

Objective: Determine research behavior of users. Determine "points of pain".

9. Are you satisfied with the number and variety of sources that you currently use? Are you confident that you are looking in all the right places?

Objective: Determine potential research needs. Determine if there is a reason to change current practices and adopt new methodologies.

10. Once you find a useful resource or article, how do you keep track of it or save it?

Objective: Determine how users manage information.

Appendix B: Assessment Plan

Purpose of the Focus Groups

The purpose of these focus groups was to document the research behaviors and needs of faculty, residents, and graduate students in the health sciences in order to inform the development of a health sciences portal. We inquired about the research habits of the participants and tried to determine the kinds of sources that they currently use.

Key Questions

The key questions addressed by this round of assessment include the following:

1. What are the research *behaviors* of users who possess medicine and health sciences expertise?
2. What are the research *needs* of users who possess medicine and health sciences expertise?

Target Audience

A health sciences portal has two potential target user groups: medical professionals involved directly in patient care and health sciences faculty and graduate students.

Subject Selection Criteria

For this set of focus groups, Rebecca Davis and Jo Anne Boorkman recruited residents, faculty, and graduate students to participate in the focus groups.

Needs Assessment Design

This round of needs assessment consisted of four focus groups: two with residents and two with faculty and graduate students.

Focus Group

Each focus group consisted of a facilitator, four to eight participants, and an observer. The facilitator used a list of prepared questions as a guide for the discussion. At the end of each faculty and graduate student session, participants were asked to fill out a short survey.

Observation

During the focus groups, an observer took notes about the discussion.

Data Collection Methodology

The observer took notes on participants' responses to the facilitator's questions, as well as any ensuing discussion using a laptop computer. In addition, the sessions were recorded using a digital voice recorder.