

## Web-scale search and virtual reference service: How Summon™ is impacting reference question complexity and reference service delivery<sup>1</sup>

William Meredith  
Royal Roads University

### Abstract

Web scale discovery tools like Summon™ are becoming the norm at academic libraries across North America. How much do these tools simplify discovery? What changes do they bring to the provision of reference service? At Royal Roads University, where most students complete graduate degrees through distance study, I applied the READ scale to reference questions received by email in the year before and the two years after we adopted Summon™. Comparing the questions over time, and analyzing reference statistics, I reflect on changes brought on by discovery layers and what they mean for the future.

**Keywords:** discovery layers, summon, reference service, digital reference, reference question complexity

### Introduction

Web scale discovery tools such as Summon™ by Serials Solutions, Primo by Ex Libris, and EBSCO Discovery Service by EBSCO are becoming the norm at American and Canadian university libraries. Breeding's market survey (Breeding, 2012) suggest that the tipping point for discovery layer adoption occurred in 2009.

The value proposition of a single search box through which our users can search nearly all of our textual material is too promising to ignore. Lippincott (2005), Mizrachi (2010) and Head and Eisenberg (Head & Eisenberg, 2009) found that students typically begin their academic research on the open web rather than on library homepages or within the databases licensed on their behalf. With competition from open web tools like Google and Google Scholar, university libraries are justified in considering discovery layers to be essential in leading students into the library's licensed online resources.

---

<sup>1</sup> This is an electronic version of an article published in 'Internet Reference Services Quarterly', Volume 18, Issue 1, pages 1-13. Internet Reference Services is available online at <http://www.tandfonline.com/loi/wirs20#.UzNC-k1OWUk>

Discovery layers are sold with the promise of simplifying search. The expectation is that students and faculty will more easily find the material they need. The corresponding implication is that demand for reference services will diminish. However, at Royal Roads, implementing Summon™ had the opposite effect. In the twelve months after implementing Summon™, we received 29% more reference questions. While our full-time student enrollment also climbed by almost 10% in that period, there were no other changes in service delivery, instruction practices, or library marketing to explain this disproportionate increase.

Perplexed by this spike in demand, curious about the impact Summon™ was having on library users' research, and interested to consider what Summon™ might mean for reference service delivery into the future, I began examining how reference questions and service delivery had changed so far. I wanted to know if the single search box led to more sophisticated questions. My hypothesis was that it might be prompting researchers to shift from asking *where* to search to asking *how* to search? I chose to analyze reference interactions conducted via email since these interactions are captured completely and are most open to objective scrutiny.

### **The setting**

Royal Roads is a small university on the outskirts of Victoria, British Columbia, Canada. The university specializes in delivering applied programs through a blended model of distance learning and short on-campus residencies; about two thirds of programs are delivered at the graduate level.

Of the reference transactions completed during the period of this study, 76% were handled by email, 11% were handled by phone, and 13% were handled in person. Email reference service is guided by a "Best Practices" document described by Croft and Eichenlaub (2006). Emailed requests for reference help are sent to a shared address of [reflib@royalroads.ca](mailto:reflib@royalroads.ca). Available librarians claim given questions and the combined interactions are stored in a shared email directory. The Royal Roads librarians have always taken pride in the quality of research support they offer. Students and faculty receive customized, detailed, same-day replies. Survey feedback on our email service has always been positive.

All students enrolled in degree granting programs attend in-class library instruction sessions, and are therefore exposed to the suite of library resources at their disposal. These sessions are typically mandatory, last about ninety minutes, and enable librarians to explain library services, demonstrate library resources, and be seen as friendly and available to help. In the first two years covered by this study, these sessions were offered during both the first and second on campus residencies. In the third

year of this study, second residency instruction was replaced by librarians embedding within the webpages of select research-intensive online courses.

### **Implementing Summon™ and the initial uptake**

The decision to implement Summon™ at Royal Roads was made well in advance of the service going live. We began describing Summon™ in library instruction sessions as much as a year in advance of making it available so that students would understand the tool when it was launched. On February 10<sup>th</sup>, 2010, we unveiled Summon™ along with a broad re-design of the Library's home page. We moved away from a website with three columns of links, the centre column including links such as 'Search for Articles' 'Search for Books' and 'Search RRU Journal List'. The new homepage contained a tabbed search box prominently displayed in the upper middle section with four blocks of links beneath. The default tab was set to Summon™ with the other three tabs offering quick access to 'Books', 'Journals by Title' and 'Article Databases'.

There is reliable evidence that students began using Summon™ immediately. In the fall of 2011, eight months after launching the new library website, two of my colleagues (Croft & Mussell, 2012) conducted a survey of library users. Respondents were asked to identify which research tools they had used to date and report on the research path they followed in their most recent assignment. 66% of students had used Summon™ before the survey, with 22% declaring it their first choice of research tool and 24% declaring it as their second choice. The 66% use-rate for Summon™, though, must be considered with the knowledge that as many as twenty percent of respondents were brand new students who had not yet begun their programs or used any library resources. Thirty nine percent of respondents defined their results from Summon™ as "essential", as compared with 35% percent of results from Google Scholar and 64% of results from subject specific databases. More than 65% of respondents declared that Summon™ had improved their ability to research effectively.

So we know that a significant number of students at Royal Roads were using Summon™ and that they were largely satisfied with both its ease of use and its results. My goal was to build on Croft and Mussell's (2012) findings and to quantify how reference questions had changed since the implementation of Summon™ .

### **Literature review**

The evolution of categorizing reference questions began with classifying them by type rather than by complexity. Christ (1947, p. 25) provided familiar categories of 'directional', 'instructional', 'advising'

'ready reference', 'search' and 'government documents'. Katz (2002, p. 16) provided a similar perspective, but with four categories rather than six: 'directional', 'ready reference', 'specific search' (which includes information on a specific topic) and 'research', (which is described as a detailed question perhaps requiring a specialist). Classifying questions by type has been used more recently as well; investigating if and how reference questions changed with the emergence of digital reference services, Schwartz (2004, p. 9), Diamond and Pease (2001, p. 214), Fennewald (2006, p. 25), and Duff and Johnson (2001, p. 49) all developed their own similar but unique scales against which to measure question type.

Other researchers have included subject categories in codifying reference questions. Brown (1985, p. 295) used subjects to compare questions received over the phone against questions received in-person. Brown also coded the sources required to meet the user's need, namely the librarian's personal knowledge, the card catalogue, ready reference sources, general reference sources and the circulating collection. Carter and Janes (2000, p. 255) applied subject categories to questions received by the Internet Public Library in 1999.

The two classification systems I considered using were the Warner Model (Warner, 2001) and the READ scale (Gerlich and Berard, 2007). Both of these scales were designed to capture the relative complexity of responses required, positioning them along a continuum. Thus, they were both better suited to my needs than the options listed above. To determine which of these scales best suited my need, I applied them both to a non-randomized sample of 75 recently received questions.

The Warner system (2001) was developed as a tool to triage incoming reference questions, routing the more complex questions to the librarians and the less complex questions to support staff. The system provides four categories:

1. 'non-resource based', which would include directional or circulation questions
2. 'skill-based', which would require a demonstration but could be replaced with clearly written directions
3. 'strategy-based' which requires an individualized approach
4. 'consultation' which implies detailed and collaborative work with a researcher or instructor.

(p.53)

The READ scale (Gerlich and Berard, 2007) is more nuanced than the Warner system because it uses six categories rather than four. While Gerlich and Berard themselves found that this increased nuance led to discrepancies in coding practices between librarians, the possibility of discrepancies did not concern

me much since I planned to do all the coding myself. What really attracted me to the READ scale, though, beyond the nuance provided by six categories, were the category labels which I found to be detailed and the most clearly suggestive of rising complexity. The full description of the READ scale can be reviewed at [readscale.org/](http://readscale.org/) but for the purposes of my analysis I distilled the labels down to the following:

1. Least effort/referral
2. Known items/call numbers/library policy
3. Direction to DBs/basic search help/technical problems
4. Complex search strings/refining topic/citation mining
5. Cooperative with researcher
6. In depth/working for researcher

### **Examining reference statistics captured in-house**

The Royal Roads librarians use a detailed online form to collect reference statistics. We track the date, time of day the question arrived, how many hours the user waited for a reply, the medium of the transaction (email, telephone, or in person), the time it took to craft the response, the user's program of study, and which librarian delivered the reply. There is also a notes field and a list of check boxes for common types of questions such as known item requests. Results are captured in an Excel<sup>®</sup> spreadsheet for analysis.

The date of given reference interactions, and the medium of the exchange, were fundamental to my research as they helped me determine how many questions fell in or out of the scope of my analysis. Beyond that, the most useful value from these statistics was time spent on responses. The time librarians` spent answering given questions is a useful measure of how reference service delivery changed with the launch of Summon<sup>™</sup>.

### **Methodology**

All the reference questions analyzed in this study were either coded or reviewed in March 2013 to ensure consistent coding practices. To code them, I created three folders within the librarians' shared Microsoft Outlook<sup>®</sup> email directory, one for each 12 month period being studied. I copied the contents of these folders into Access<sup>®</sup> database software since the software made the process fairly easy. I then imported the Access<sup>®</sup> files into Excel<sup>®</sup> spreadsheet software because this software made it easier to

randomize the rows. I inserted a blank column into each Excel® sheet and populated it with numbers randomly generated by the software. I sorted the sheets according to these random numbers and imported the files back to Access®. I then created a form in Access® to sort through the emails one by one, coding each for complexity. This form listed select fields such as the name of the sender, the date of the interaction and the content of the exchange. I added a dropdown menu to capture either the code for complexity or the "not applicable" category used to disqualify emails sent between colleagues or messages that followed up on a phone call or were unsolicited follow-up replies to earlier questions.

In selecting my sample size of reference questions to code, I chose a 95% confidence level with a 4% margin of error. In 2010-11 year of this study, we responded to 1861 emailed questions so I coded 454 questions. In 2011-12 we responded to 2874 emailed questions so I coded 497 questions. In 2012-13, we responded to 2443 emailed questions so I coded 482 questions. Results are presented as the proportion of questions received annually rather than as real numbers to make the annual totals comparable.

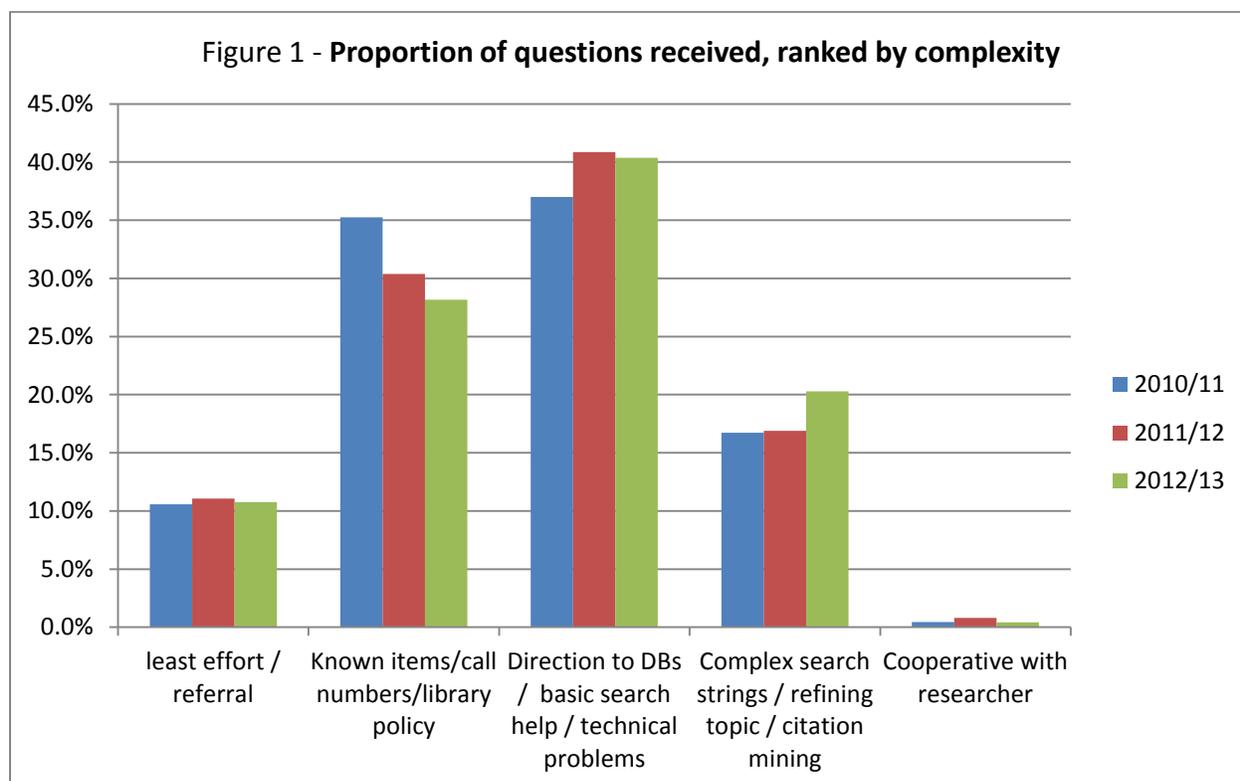
The biggest challenge to coding question complexity was assigning values to 'known item' requests. For these questions, the answer we provided often dictated the value I assigned. Known items were coded as category two unless an explanation of how to find the item was required. If someone failed to find something that was in our collection or available on the open web, a complete answer required that we explain how they could have found it for themselves. Such interactions were coded as a level three because they required directions to a database and some basic searching. Coding these interactions as level three was also appropriate because sometimes people could not find a document because of a technical issue, and level three also made an allowance for that.

I applied category four (complex search strings/refining topic/citation mining) to questions requiring a detailed response, generally including an exploration of search terms and an explanation of how to develop complex searches with Boolean operators.

The "cooperative with researcher" category is underrepresented in my results since every email interaction had to be considered on its own. While a complex research need may have been met over the course of a few emails, different components of the overall interaction may have required different levels of service so it was necessary to code them separately. Also, because I was coding a random sample of questions, coding the conversation rather than the specific interaction would have led to duplicate counts for the same initial question.

## Results

Figure 1 shows that Summon™ had little impact on the distribution of complex questions we received. All but one of the changes I found fell within my margin of error. Users may have appreciated the single search box but this study provides questionable evidence that their search strategies changed as a result.



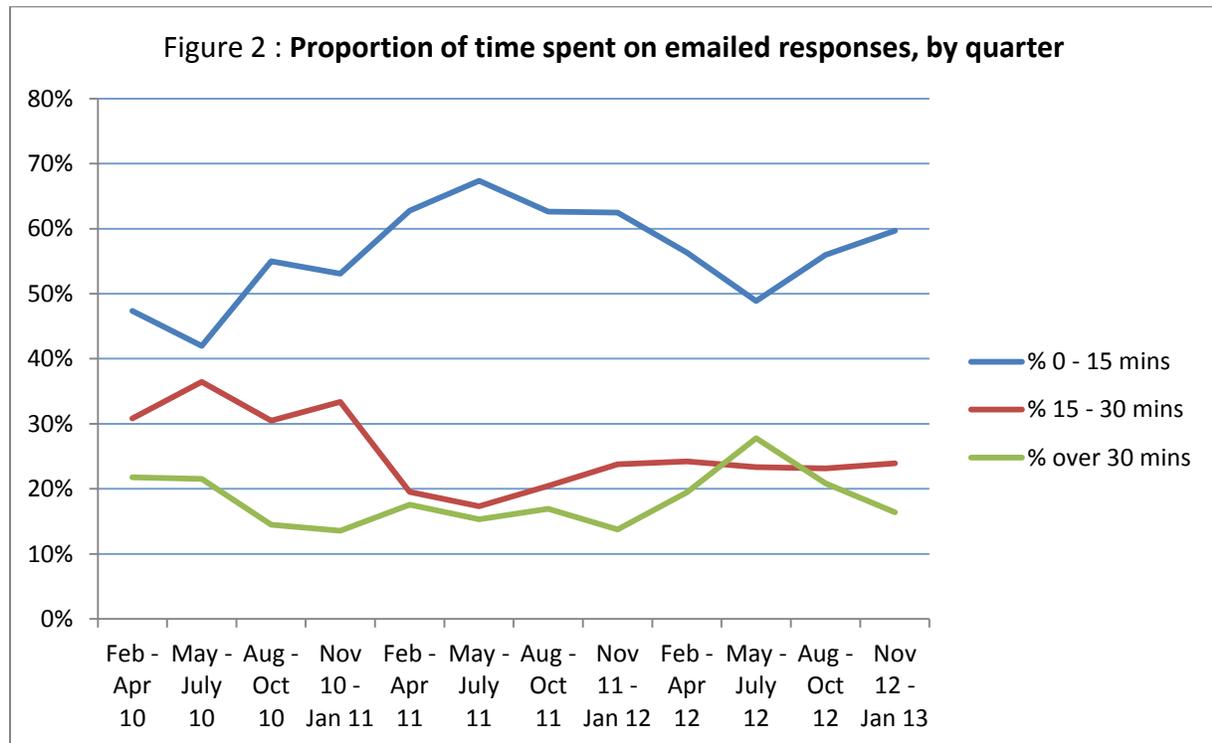
**Figure 1:** Proportion of questions received, ranked by complexity

The greatest change was seen in the proportion of category two questions (known items/call numbers/library policy). These fell by 4.8% in the year after we launched Summon™ and a further 2.2% in the year after that. Summon's™ single search box provides the most simple method possible to seek out known items so the drop in this category is unsurprising.

It was in the more complex categories that I was curious to measure change. I wondered what shift might have occurred between category three (direction to DBs/basic search help/technical problems) and category four (complex search strings/refining topic/citation mining). My hypothesis had been that the Summon™ might be leading people to shift from asking where to search to asking how to search more effectively. This study demonstrated that the introduction of Summon™ brought about little

change. In the year after implementing Summon™, the proportion of category three questions rose by 3.8% and then dropped by .4% in the third year. The proportion of complex search questions rose by .2% in year after we launched Summon and a further 3.4% in the third year.

Although Summon™ had little impact on the nature of the questions we received, it did influence the librarians' reference service. Figure two shows how the time we spent answering questions shifted. In the year after adopting Summon™, the proportion of questions that took us fewer than fifteen minutes to reply to rose by more than 14%, although it then dropped back to 5.7% above the pre-Summon™ year. The proportion of questions that took between fifteen and thirty minutes to respond to dropped by 12% the first year and only rose slightly after that. The proportion of questions that took more than thirty minutes to respond to showed no enduring change, dropping initially but climbing to slightly above pre-Summon™ levels in year 3.



**Figure 2:** Proportion of time spent on emailed responses, by quarter

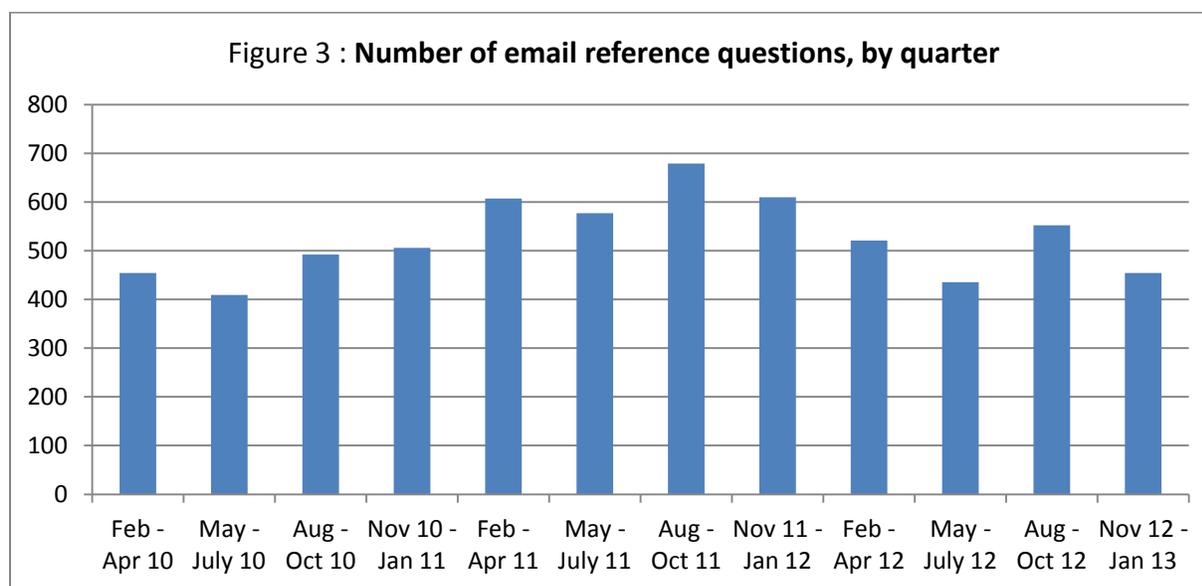
### Limitations

It is worth considering if the READ scale was the appropriate instrument for this study. I believe that it was and that the categories it includes are clear and distinct. Category three, (direction to DBs/basic search help/technical problems) collocates 'technical problems' with the very different 'direction to

databases', but the intent of the READ scale is to provide a continuum of complexity so placing these two kinds of questions together makes ample sense. Having these two elements in the same category also worked well for some known item searches. When someone neglected to find an item that was in our collection, the appropriate response was to provide directions to that item in its native database. However the reason a person was not able to find an item was frequently related technical hurdles, such as faulty indexing or link resolver failure. Breaking category three into two distinct values would complicate the coding process and make the results less reliable.

### Discussion

Demand for reference help rose substantially in the twelve months after we launched Summon™ and, even taking into account the rise in FTE student enrollment, remains higher now than it was before we adopted the discovery layer. But beyond a diminishing demand for help locating known items, my findings illustrate that Summon™ has not had much of a positive nor a negative impact on the sophistication of the questions we received. This lack of change was disappointing but not a complete surprise. Meadow and Meadow (2012) mined Summon™ search logs and also found no improvement over time in researchers' approach.



**Figure 3:** Number of email reference questions, by quarter

One of the benefits of Summon™ is that the single search box allowed us to be quicker in our replies. As I mentioned earlier, and as is illustrated by figure 3, the number of reference questions we received rose by 29% in the twelve months after adopting Summon™. It seems likely that the timing of this spike in

demand was a coincidence, but having the Summon™ search box in the middle of our homepage enabled us to reply faster by freeing us from explaining how to navigate to whatever specific database we might have referred people to before. The tool also gave us new capacity to repurpose previous answers. Responses to complex questions about organizational management, for example, which explored possible search terms and explained complex search syntax, could be repurposed and customized to use in answering a cultural studies question.

Question complexity is not increasing because students and faculty remain as unsure as they ever were about how to conduct library research. Giving them a single search box does not change their level of confusion about the search process. Meadow and Meadow's (2012) analysis of Summon™ search logs found that the most frequent Summon™ search entered was "facebook" and that most subject searches consisted of a few keywords. Fagan et al. (2012), as part of their usability study of the EBSCO Discovery Service, assigned participants a known item search and provided them with a subject database recommendation; three of ten participants chose to skip this question and six began within EDS anyway, searching for either the database name or the known item. Users' confusion extends to processing results as well. Thirty percent of Fagan et al's respondents reported feeling overwhelmed by their result sets but could not explain when not to use the tool. Gross and Sheridan's usability study of Summon™ found that users had trouble differentiating results, for example distinguishing books from book reviews (Gross & Sheridan, 2011, p. 242).

Head and Eisenberg's (2009) study of the research practices of 2,318 college students on six campuses across the United States shed important light on students' research practices. They found that students are pragmatic and pressed for time. Once students find a successful approach to searching for literature, they follow that same approach for all their academic research regardless of discipline. While the researchers found that students typically begin with Google or Google Scholar, they also found that 84% then move on to also use scholarly databases. Head and Eisenberg collected their data before the widespread introduction of discovery layers like Summon™, but research strongly suggests that students with access to discovery layers are now turning to them once they are finished with Google (Croft & Mussell, 2012; Fagan, Mandernach, Nelson, & Paulo, 2012; Gross & Sheridan, 2011). The new challenge to providing good quality reference support, then, is in helping students understand the scope and merits of discovery layers and educating about specific subject specific databases that might add value to what they have already found.

The power of the single search box has both positive and negative connotations for reference service. The most obvious benefit of these tools is that they do what they promise: simplify discovery. Many students and even faculty can meet their information needs by plugging some keywords into the search box and filtering their results. Another benefit is that it can simplify aspects of information literacy instruction. The simple search box, relevant to almost all disciplines, enables librarians to focus less on interface specifics and more on the process of effective search: developing search terms, understanding Boolean operators, constructing nested searches, filtering results, evaluating material and building on those results. This is useful for the students and faculty, helping them to understand how they might search more efficiently, as well as for the librarians, enabling them to provide service faster because they do not have to explain multiple sources and variations in functionality. The central negative implication of the single search box is that we may be encouraging students to imprint on a less useful tool than what they need and disregard the value of looking elsewhere.

When students come to us with questions, we need to be ready to explain what tools are best for their needs. When Croft and Mussell (2012) surveyed Royal Roads library users after we launched Summon™, and asked them to rate their satisfaction with results from different online resources, 64% identified results from subject specific databases as essential compared with 39% from Summon™ and 35.4% with Google and Google Scholar. While Summon™ is obviously a valuable tool, librarians must be careful not to lapse into neglecting subject specific resources able to return more precise results.

### **Further research**

My research suggests that the type of support students and faculty require from librarians did not change much with the launch of Summon™, but that is not to say that student and faculty research processes have not been changed by Summon™ or that users have not developed opinions about the supposed ease of searching or the relevance of their result sets. Rather, we need different ways to measure the changes.

Further research could shift attention away from the questions asked and onto the answers given. How have responses to reference questions changed with the advent of discovery layers? One simple measure could be counting the resources suggested in each librarians' response. In coding the more than fourteen hundred responses considered for this study, I sensed an increasing reliance on Summon™ and a decreasing reliance on other tools perhaps better suited to the researchers' needs.

The best measure of the worth of a discovery layer would be one that quantifies changes in students' grades, but this of course presents challenges in methodology. In order for results to be comparable over years, this line of inquiry would require working with instructors who had given the same assignment, and used the same grading rubric - a rubric that assigns value to the quality of the sources used - before and after Summon's™ launch.

Comparing database use statistics before and after Summon™ is another way to gauge changes in users' research practices, as well as the library's return on investment for spending on electronic resources. Way's (2010) examination of database use at Grand Valley State University before and after implementing Summon™ showed that click throughs from aggregator and publisher databases dropped almost across the board while the number of full text downloads climbed noticeably. These changes demonstrated that users were increasingly doing their searching through Summon™ and finding more relevant material than they did before. It seems likely that Way's results would be duplicated if the same measures were applied at other libraries.

The limitation of the areas of further research listed above is that they do not measure users' satisfaction with the search process or their results. To determine how satisfied people are with Summon™, we have to ask them directly. Inspiration for doing so can be found in Gross and Sheridan's (2011) usability study of Summon™. In the first semester of 2010 they asked five novice undergraduate researchers to find answers to four research questions. Similarly, Fagan et al. (2012) conducted a usability study of EBSCO's EDS in October of 2010, asking a two faculty and eight students, with different research experience, to find answers to nine research questions.

### **Conclusions**

The librarians' value as information stewards remains secure, and may in fact be more solidly rooted than it was. Discovery layers promise simplified discovery and imply a trend towards disintermediation between users and the research they need. As ranking algorithms and interface designs improve it seems inevitable that these tools will come ever closer to fulfilling that promise. However, at least for the time being, Summon™ is another interface that users need assistance navigating, and looking beyond.

One could argue that as valuable as it to have a discovery layer such as Summon™ these tools have in fact made the provision of good reference service more complicated rather than less. In addition to guiding users through the subject specific databases that information professionals have relied on in the

past decade or so, we must now also support users in navigating and zeroing in on research topics using these broad-based tools. With the simplification of search promised by discovery layers, it is all the more important that we lead researchers to the most important content available. But it is imperative that we meet this challenge and at least educate users on the relative merits of the range of tools. Head and Eisenberg (Head & Eisenberg, 2009) found evidence that 84% of college students did use scholarly databases in doing their research. They also found that only 10% sought research help from librarians. If we fail to impress the minority of students and faculty who come to us for support, and who may have already tried Summon, we risk losing their attention in the long term and being seen as irrelevant.

## Reference List

- Breeding, M. (2012). Automation Marketplace 2012: The Complete Survey Data. *The Digital Shift*.  
Retrieved from <http://www.thedigitalshift.com/2012/03/ils/automation-marketplace-2012-the-complete-survey-data/>
- Brown, D. M. (1985). Telephone Reference Questions: A Characterization by Subject, Answer Format, and Level of Complexity. *RQ*, 24 (3), 290–303.
- Carter, D. S., & Janes, J. (2000). Unobtrusive Data Analysis of Digital Reference Questions and Service at the Internet Public Library: An Exploratory Study. *Library Trends*, 49 (2), 251–265.
- Christ, R. W. (1947). Recording Reference Service. *College and Research Libraries*, 8 (1), 23.
- Croft, R., & Eichenlaub, N. (2006). E-mail Reference in a Distributed Learning Environment. *Journal of Library Administration*, 45 (1-2), 117–147. doi:10.1300/J111v45n01\_07
- Croft, R., & Mussell, J. (2012). Discovery Layers and the Distance Student: Online Search Habits of Students. *Journal of Library & Information Services in Distance Learning*, 7 (1-2), 18–39.
- Diamond, W., & Pease, B. (2001). Digital Reference: A Case Study of Question Types in an Academic Library. *Reference Services Review*, 29 (3), 210.
- Duff, W. M., & Johnson, C. A. (2001). A Virtual Expression of Need: An Analysis of E-mail Reference Questions. *The American Archivist*, 64 (1), 43–60.
- Fagan, J. C., Mandernach, M., Nelson, C. S., & Paulo, J. R. (2012). Usability Test Results for a Discovery Tool in an Academic Library. *Information Technology and Libraries*, 31 (1), 83.
- Fennwald, J. (2006). Same Questions, Different Venue. *The Reference Librarian*, 46 (95), 21–35.  
doi:10.1300/J120v46n95\_03
- Gross, J., & Sheridan, L. (2011). Web Scale Discovery: The User Experience. *New Library World*, 112 (5/6), 236–247. doi:10.1108/03074801111136275

Head, A. J., & Eisenberg, M. (2009). *Lessons Learned: How College Students Seek Information in the Digital Age* (p. 42). The Information School, University of Washington. Retrieved from

[http://projectinfolit.org/pdfs/PIL\\_Fall2009\\_finalv\\_YR1\\_12\\_2009v2.pdf](http://projectinfolit.org/pdfs/PIL_Fall2009_finalv_YR1_12_2009v2.pdf)

Lippincott, J. K. (2005). Net Generation Students and Libraries. *Educause Review*, 40 (2), 56.

Mizrachi, D. (2010). Undergraduates' Academic Information and Library Behaviors: Preliminary Results.

*Reference Services Review*, 38 (4), 571–580. doi:10.1108/00907321011090737

Schwartz, J. (2004). Toward a Typology of E-Mail Reference Questions. *Internet Reference Services*

*Quarterly*, 8 (3), 1–15. doi:10.1300/J136v08n03\_01

Warner, D. G. (2001). A New Classification for Reference Statistics. *Reference & User Services Quarterly*,

41 (1), 51–55.

William Katz. (2002). *Introduction to Reference Work* (8th ed.). Boston: McGraw-Hill.