

2016

Territory Folks Should all be Pals: Qualitative Use of Search Logs to Improve Confidence in and Communication about a Library Discovery Service

Jimmy Ghaphery

Virginia Commonwealth University, jghapher@vcu.edu

Mark Baggett

University of Tennessee, Knoxville, mbagget1@utk.edu

Thomas McNulty

Virginia Commonwealth University, tmcnulty@vcu.edu

Emily Owens

Virginia Commonwealth University, eaowens@vcu.edu

Follow this and additional works at: http://scholarscompass.vcu.edu/libraries_present

 Part of the [Library and Information Science Commons](#)


Downloaded from

http://scholarscompass.vcu.edu/libraries_present/49


This Presentation is brought to you for free and open access by the VCU Libraries at VCU Scholars Compass. It has been accepted for inclusion in VCU Libraries Faculty and Staff Presentations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.



Territory Folks Should all be Pals



Qualitative Use of Search Logs to
Improve Confidence in and
Communication about a Library
Discovery Service



Questions about Primo

Do perceptions of search relevance vary by department?

Can conversations about these perceptions improve inter-departmental relationships?

Can this all lead to a better discovery service?

The Project

Random selection of anonymized Primo user queries

Library staff from different departments repeat searches and rate relevancy

Scores are compared and discussed

Replicate at a second research library

Two Universities

	University of Tennessee (UT)	Virginia Commonwealth University (VCU)
Enrollment	26,737 (6,400 FT Grad)	26,106 (4,984 FT Grad)
Research	\$261 million	\$270 million
Doctor's Degrees	1,141	329
Carnegie Profile	Public Research University (Very High) Balanced arts & sciences/professions, high graduate coexistence Comprehensive programs, with medical/veterinary school	

Two Libraries

	UT	VCU
ARL	Yes	No
Titles Held	2,674,798	2,543,274
Library Expenditures	\$27 million	\$21 million
Staff	248 (103 professional)	179 (64 professional)
Gatecount	2,087,299	2,058,082
Initial Circ Transactions	150,528	62,154
Database Searches	4.6 million	4.1 million

Source: ASERL Fall 2105 Statistical Index

Two Library Systems

	UT	VCU
Alma	2014	2012
Primo	2009	2011
Search approach	Tabbed	All in one

The Project @ VCU | Context

Fall 2013 Library Confidence in Primo Low

Instruction Librarians stranded in classes when system under load

Performance issues addressed, concerns remained on relevancy

Spring 2014 Relevancy study to both improve Primo and interdepartmental relations

4 Systems Librarians | 3 Instruction Librarians

The Project @ VCU | The Plan

Selected and anonymized search queries from “typical week” October 2013

Started with 200 queries for test run, scaled back to 100

Refined query selection and relevancy scale

The Project @ VCU | The Scale Part 1

- 1: Very Poor (no relevant items, misleading results, known item not findable)
- 2: Below Average (few relevant items, known item buried)
- 3: Average (some relevant items; known item on first page or accessible from reasonable use of facets)
- 4: Above Average (topic search returned mostly relevant results, known item on top 5 results)
- 5: Excellent (topic search returned very relevant results; known item search retrieved item in top 2 results)

The Project @ VCU | The Scale Refined

Please indicate how strongly you agree with the following statement: Primo produced relevant results for this search:

1 = Strongly Disagree

2 = Disagree

3 = Neither Agree nor Disagree

4 = Agree

5 = Strongly Agree

Flag for follow-up: Yes/No

Comments (optional)

The Project @ VCU | Results

Results shared prior to followup meeting

Departments	Average	Individuals
Systems	3.57	3.27, 3.56, 3.54, 3.94
Teaching	3.76	3.77, 3.66, 3.86

The Project @ VCU | Takeaways

After meeting to review findings and discuss specific examples flagged for followup

1. Forwarded problematic searches to Ex Libris
2. Several librarians report Google Scholar or Summon benchmarking
3. Pursued local customizations including “AdWords”
4. Qualitative improvement in search perception and collegiality

The Project @ VCU | Round Two

Spring/Summer 2015, another round of ratings with previous search terms

Ex Libris launched a significant improvement in the relevancy ranking algorithm

4 Systems Librarians, 2 Teaching Librarians (repeats)

Outreach Liaison Librarian, Digitization Librarian (new)

Eager to see differences (system performance or opinions)

The Project @ VCU | Round Two Results

Departments	Average Round 1	Average Round 2
Systems	3.57	3.84
Teaching	3.76	3.96
Other Librarians (Outreach & Digitization)		3.48
Search Type		
Known Item	3.9	4.24

The Project @ VCU | Round Two Discussion

Summary results not shared before discussion

Systems librarian pulled samples for discussion (many comments or large variations in ratings)

Perceptions

- Performance of the discovery tool was better (for those repeating)
- Biggest gap in broad topic searches
- Known item searching ok
- Appreciation for the project

The Project @ VCU | Sample Searches

Is it too much to ask to anticipate these types of searches:

- [1960s School Segregation](#) (want 1960s or 1960's etc)
- [Proc Natl Acad Sci USA 69:907-911](#) vs [Google Scholar](#) (citation search)

Or to do something better with a broad search like [journalism](#)

Library search is hard. Good or bad result? [Tennis shoes](#)

Round two unanticipated local adwords boost: [scifinder](#)

The Project @ Tennessee | Context

2009 - Implemented Primo

- “Out-of-the-box” Approach
- Departmental upheaval left us without a *true* Primo administrator

2013 - Discovery & Access Working Group (DAWG)

- Improve discovery interface
- Implement Primo Central
- Anticipate that the Aleph Web OPAC would eventually go away

2016 - Perceptions of Primo as a Discovery Tool

- Understand the differences in perception of Primo between “public” and “technical” services staff
- Conduct quantitative survey
- Perform qualitative post-survey discussion

The Project @ Tennessee | The Plan

Programmatically selected and anonymized search queries from all queries in February 2016

Started with 100 queries, scaled back to 50 (Advice of Departmental Review Committee)

Programmatically assigned strings related to each query (One Search, UT Collections, Advanced Search, Browse, *Deep Link*, *Internal Search*)

Invited all DAWG committee members (19) to participate

Asked to self identify as *public services* or *technical services*

The Project @ Tennessee | Snapshot of Logs

Type	Number of Queries	Percentage of Total
Total (February 2016)	122570	100%
One Search	108069	88%
UT Collections	13333	11%
Advanced Search	7753	6%
Browse	1168	~1%
Deep Links	42715	35%

The Project @ Tennessee | Rubric

- 1: Very Poor (no relevant items, misleading results, known item not findable)
- 2: Below Average (few relevant items, known item buried)
- 3: Average (some relevant items; known item on first page or accessible from reasonable use of facets)
- 4: Above Average (topic search returned mostly relevant results, known item on top 5 results)
- 5: Excellent (topic search returned very relevant results; known item search retrieved item in top 2 results)

Respondents were allowed to skip queries where they found the query results difficult to interpret.

The Project @ Tennessee | Survey Results

Survey results

Departments	Average	Individuals
Public Services	3.54	3.74, 2.9, 3.63, 3.49, 3.3, 4.18
Technical Services	3.9	4.18, 4.12, 3.94, 4.02, 3.22

The Project @ Tennessee | Results

Comparing Averages Across Query Types

Query Type	Public Services Avg.	Technical Services Avg.
OneSearch	3.38	3.75
UT Collections	4.1	4.24
Advanced Search	3.6	4.3

The Project @ Tennessee | Concordance

Even though we all used the same rubric, there were some issues with inter-rater reliability.

Public Services:

- **58%** of queries had at least a 2 point swing in the high and low rating in the group
- **8%** of queries had at least 1 person rate the results a 5 and at least 1 person rate the results a 1

Technical Services:

- **56%** of queries had at least a 2 point swing in the high and low rating in the group
- **4%** of queries had at least 1 person rate the results a 5 and at least 1 person rate the results a 1

The Project @ Tennessee | Skipped Queries

Survey respondents were allowed to skip queries where the results were difficult to interpret.

7 queries were skipped by respondents in public services.

2 queries were skipped by respondents in technical services.

0 queries were skipped by more than one individual regardless of group.

The Project @ Tennessee | Queries for Discussion

In order to better understand concordance issues, **eleven** queries were selected for discussion.

Six queries were ones where one person in the group rated a query a 1 and another rated it a 5.

Five questions were queries where there was more than a one point swing between the two groups on average.

The Project @ Tennessee | Facilitating Discussion

Do you feel you have a good understanding of the searcher's intention?

Are the results returned from the query satisfactory? Why? Why not?

Based on the query, do you think the results should be better?

How do you expect a discovery tool to handle a search like this?

Why do you think respondents rated the query results differently?

The Project @ Tennessee | Discussion Highlights

A known-item that we don't have

- Public Services: "Since results are returned, I'd think this would insinuate that we have this but I'm doing something wrong."
- Technical Services: "We're following the Google model, and Google always returns something."
- Public Services: "Students have similar issues when they search for articles or books we don't have, and reviews are returned."

It's not about words; it's about "aboutness."

- Both groups: "This looks like a known item search."
- Public Services: "The results returned here have the words but I question the distance between them." "The sources of some of these results make no sense." "Technical services just sees the words; they don't think about the aboutness."

Results that invoke an emotional response

- Public Services: "This is an advanced search for couples and infertility. The first result says Female Infertility. I found that offensive."
- Public Services: "There is nothing inherently wrong with the results here, but that first result invoked an emotional response and I threw the rubric out the window."
- Matthew Reidsma's "[Algorithmic Bias in Library Discovery Systems](#)"

The Project @ Tennessee | Challenges and Next Steps

Known item searches

Search quality

DAWG-wide discussion

Conclusions: UT | VCU Compare

Discovery: *Better than average!* UT 3.71 | VCU 3.62

Context matters, defining and determining relevancy

Quantitative vs Qualitative (sample size, form size, for understanding differences in perspectives)

Comparison with other search expectations

Conclusions: General Takeaways

Pinpointing potential areas for search improvement

Effective vehicle for productive conversations about discovery

Dialog, research, and replication across institutions

References

“Building Bridges with Logs: Collaborative Conversations about Discovery across Library Departments” Code4Lib Issue 32. Jimmy Ghaphery, Emily Owens, Donna Coghill, Laura Gariepy, Megan Hodge, Thomas McNulty, Erin White

GitHub repository - <https://github.com/vculibraries/primo-log-sampler>

Thanks

Mark Baggett mbagget1@utk.edu

Jimmy Ghaphery jghapher@vcu.edu

Emily Owens eaowens@vcu.edu

Thomas McNulty tmcnulty@vcu.edu

Questions and comments welcome