

*Did We Get What We Paid For?
An Institutional Assessment of User Behavior,
User Expectations, and System Performance
of Web Scale Discovery Tools*



**Doralyn Rossmann and Brian Rossmann
Montana State University
Bozeman, Montana, USA**

Overview of the Study

- ❑ Assess an implemented Web Scale Discovery (WSD) tool with regard to user behavior, system performance, and collection coverage – did we get what we paid for?
- ❑ How others might use this study:
 - Provide tools and benchmarks for future studies
 - Inform how libraries can work with vendors and users
 - Guide library's placement and look of WSD
 - Comparing WSDs against each other

Introduction to Study

- ❑ Montana State University – acquired Serials Solutions’ Summon in July 2010 with 3-year contract
- ❑ “CatSearch” displayed and promoted and prominently located
- ❑ Want more than anecdotes
- ❑ How we want to use this study:
 - Inform how patrons use CatSearch
 - How improve functionality
 - Location and promotion decisions
 - Variety of data points and information sources
 - Renew at end of contract? Keep WSD at all?

Introduction to Study, continued

- Newness of WSD tools:
 - OCLC WorldCat Local: November 2007
 - Serials Solutions Summon: July 2009
 - Ebsco EDS: January 2010
 - Ex Libris Primo Central: June 2010
- Summon Assessment Group convenes January 2012
- Variety of studies, we chose to combine approaches for more complete picture (see paper for full literature review)
- Our study consists of 4 parts and 1 more is underway (user expectations have not been covered here, but is ongoing)

Part 1: Summon Link Analysis

- H1: The majority of successful full-text links take users three or fewer clicks to reach the full-text item
- H2: Successful linking to full-text resources improved during the first two years of implementation

Link Analysis, continued

- ❑ Methodology: 26 topics from actual Summon queries identified by member of research team
- ❑ Subject-specific searches rather than known item
- ❑ Categorized into full-text and non full-text links, successful retrievals, number of clicks
- ❑ Fall 2010, fall 2011, summer 2012
- ❑ First 25 results, for a total of 650 per time

Link Analysis, continued

Table 1. Failed links to full text	
Period of study	Percentage failure rate
Fall 2010	45
Fall 2011	23
Summer 2012	27

Link Analysis, continued

Table 2 : Full text: number of clicks to reach full text (Summer 2012 study)

Failed Link	158
Successful Link	
1 click	82
2 clicks	194
3 clicks	136
4 clicks	6
5 clicks	4
6 clicks	2
Non-full text	68
Total	650

Link Analysis, possible error sources

- ❑ The Summon system with its indexing and linking technologies
- ❑ The content provider with metadata or linking technologies
- ❑ The OpenURL resolver Serials Solutions' 360 Link
- ❑ The link from the 360 Resource Manager could be inaccurate
- ❑ The item selected by MSU Library may not be part of the Library's collection
- ❑ Our original implementation may have had errors or settings may have changed – requiring our attention and monitoring of Summon messages

Part 2: Summon Transaction Log Analysis

- Initial study: April 2011
 - H1: Queries performed within Summon are of low quality
 - H2: Query quality improved during the first two semesters of implementation
- Subsequent study: April 2012

Transaction Log Analysis, continued

- ❑ Data from Summon Administration Console
- ❑ Random sample of 100 queries per month (900 queries and 1000 queries in 1st and 2nd study, respectively)
- ❑ Categorized into seven query types:
- ❑ URL, invalid, natural language, database/journal, subject, known item, Boolean operator
- ❑ Grouped into high and low quality (database/journal questionable since some of those will not come up in search results by title)

Transaction Log Analysis, continued

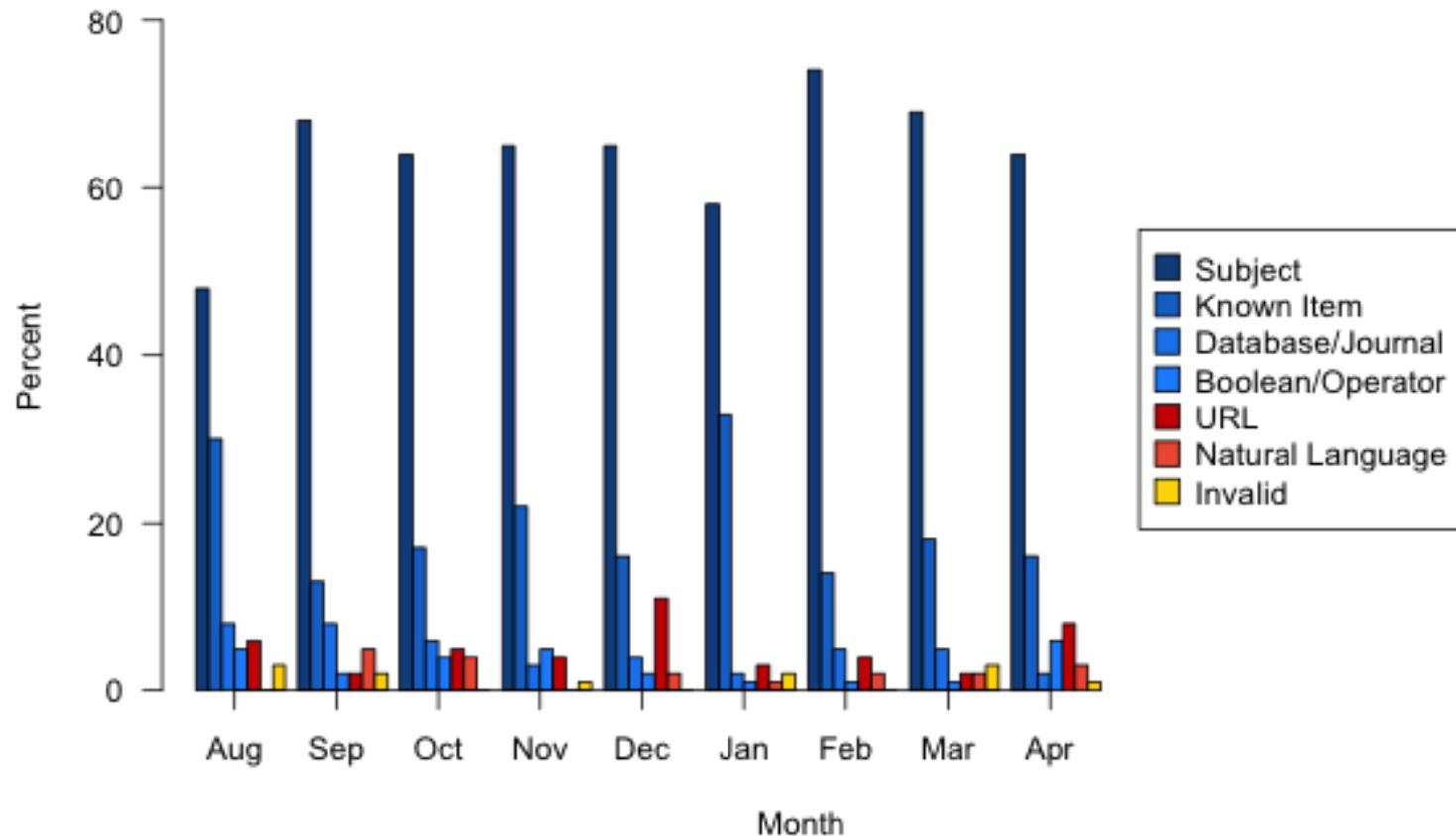


Figure 1: Bar graph displaying the frequency of query types by month for August 2010 – April 2011. Low quality queries are displayed in warm colors. High quality queries are displayed in cool colors.

Transaction Log Analysis, continued

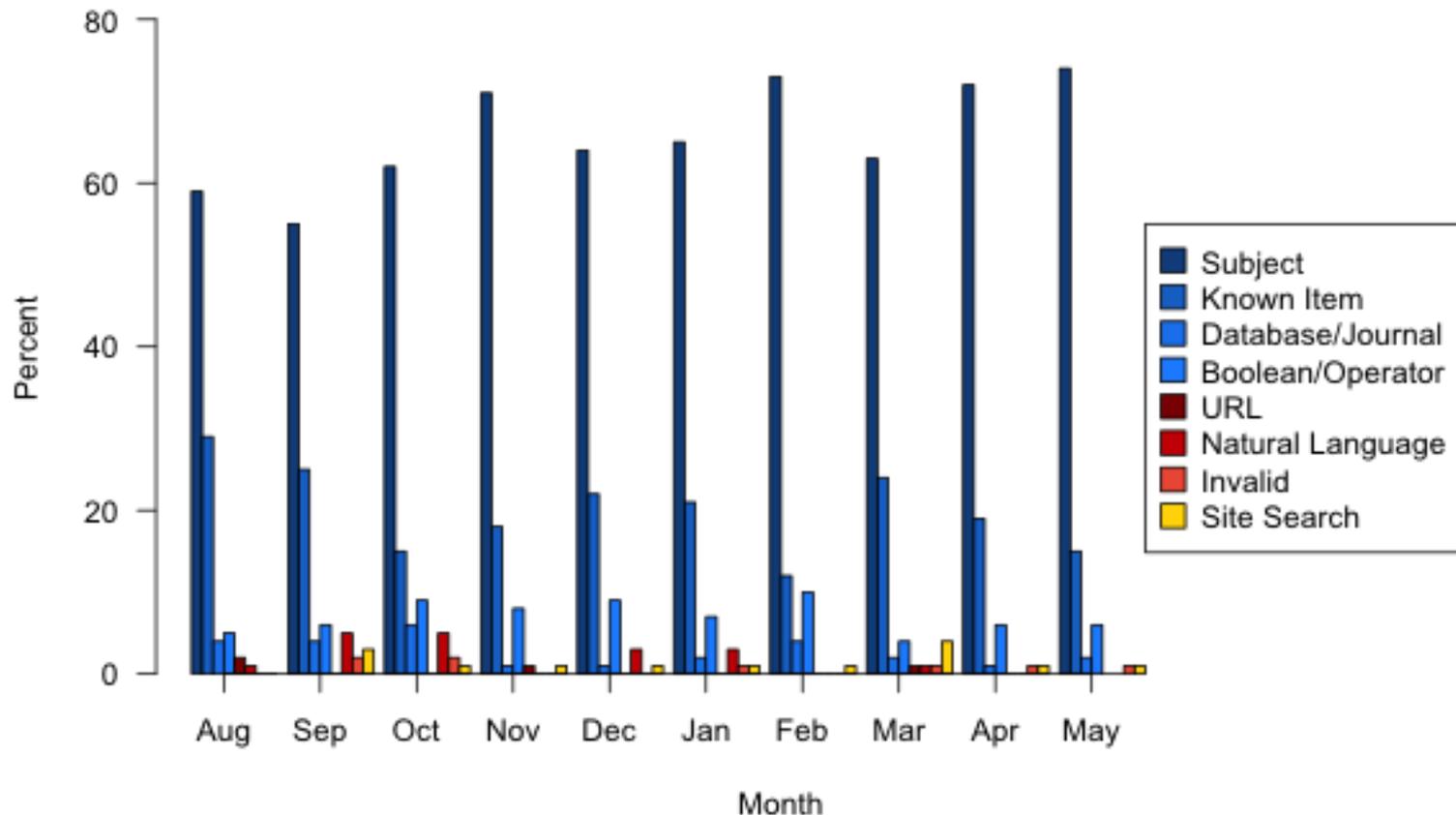


Figure 2: Bar graph displaying the frequency of query types by month for August 2011 – May 2012. Low quality queries are displayed in warm colors. High quality queries are displayed in cool colors.

Transaction Log Analysis, continued

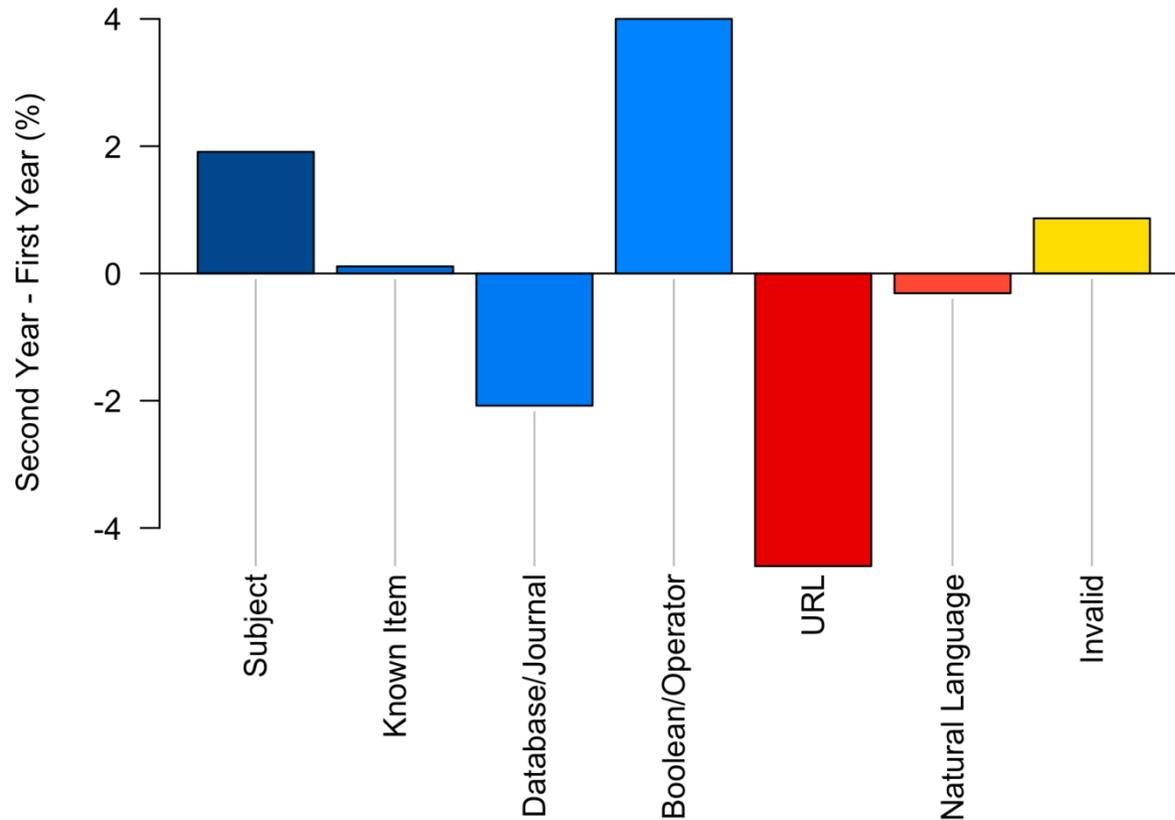


Figure 3: Percentage of each query type for each year of the study and the percent difference of each query type by year

Transaction Log Analysis, continued

Table 3: Percentage of query types for high-quality queries performed each year

Query Type	First Year (%)	Second Year (%)	Overall (%)
Subject	69.8	68.9	69.3
Known Item	21.7	20.9	21.3
Database/Journal	5.2	2.8	3.9
Boolean/Operator	3.3	7.3	5.4

Transaction Log Analysis, continued

- URLs – oddities (email with Andrew Nagy from Serials Solutions)
- New vs. known tool
- Instruction/integration efforts
- Summon data logs vs. Google Analytics

Part 3: Google Analytics (GA)

Transaction Log Analysis

- ❑ GA provides different insights into Summon use
- ❑ Understand how users navigate library web site and Summon
- ❑ Two main areas of focus: Landing Pages and Total Pageviews
- ❑ Landing pages: which page within a web site where user begins navigation (need independent GA account for Summon)
- ❑ Total Pageviews: Library web site vs. Summon

GA Transaction Log Analysis, continued

The screenshot shows the Montana State University Library website. The browser window title is "Montana State University Library - Mozilla Firefox". The address bar shows "www.lib.montana.edu". The website header includes the Montana State University logo and the text "MONTANA STATE UNIVERSITY LIBRARY". A navigation bar contains links for "MSU Libraries", "MSU Libraries Intranet", "MSU Exchange Outlook", "MSU Library Staff Intra...", and "http://twitter.com/". A search icon and "Ask A Librarian" link are also present.

The main content area features the tagline "Inspiration, Discovery, Knowledge" above a large photograph of a modern library building. Below the photo is a search bar labeled "CatSearch" with the text "Search for articles, journals, books and more" and a "SEARCH" button. Underneath the search bar is a link to "Advanced Search".

The website is organized into three main columns:

- FIND**: Find research materials, including articles, books, databases, journals, and course reserves.
- REQUEST**: Request resources and services, including group study rooms, laptops, documents, and books.
- INTERACT**: Learn about the library and meet with us for research assistance, writing help, and tech support.

On the right side, there is a sidebar with tabs for "Events", "Hours", and "Twitter". The "Events" tab is active, showing a list of events:

- 10/30 Library Workshop: Excel 2010: Beyond the Basics
- 10/30 Library Workshop: Citing Sources
- Database trial: IBISWorld Industry Research Reports
- A New Library Website

At the bottom of the page, there are social media icons for Twitter, Facebook, and YouTube. The footer contains the following text: "© MSU About MSU Library Accessibility Contact Us Privacy Policy Mobile Site Help Site Index & Site Search".

GA Transaction Log Analysis, continued

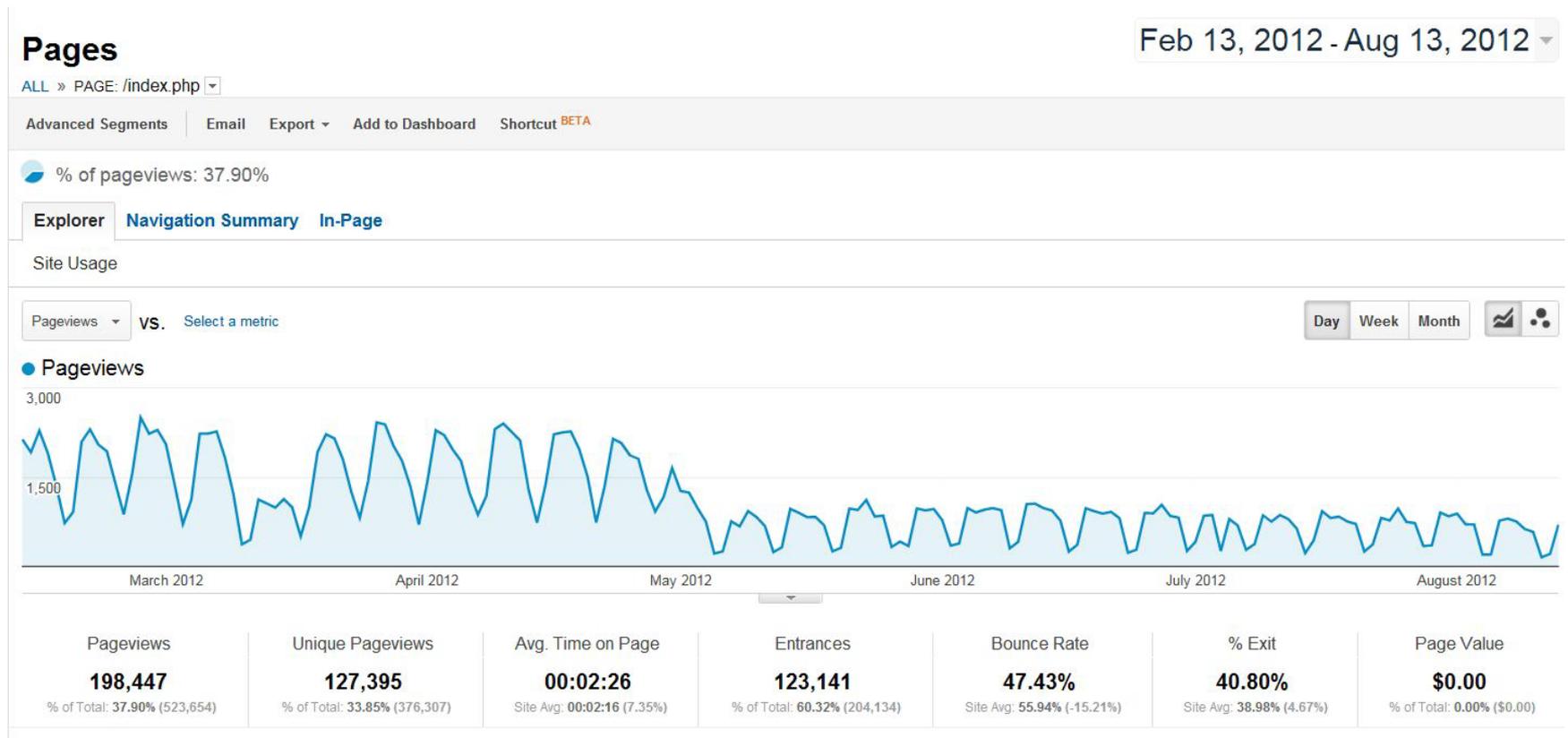


Figure 4: Total pageviews for MSU Library Summon searches in six months, 2012.

GA Transaction Log Analysis, continued

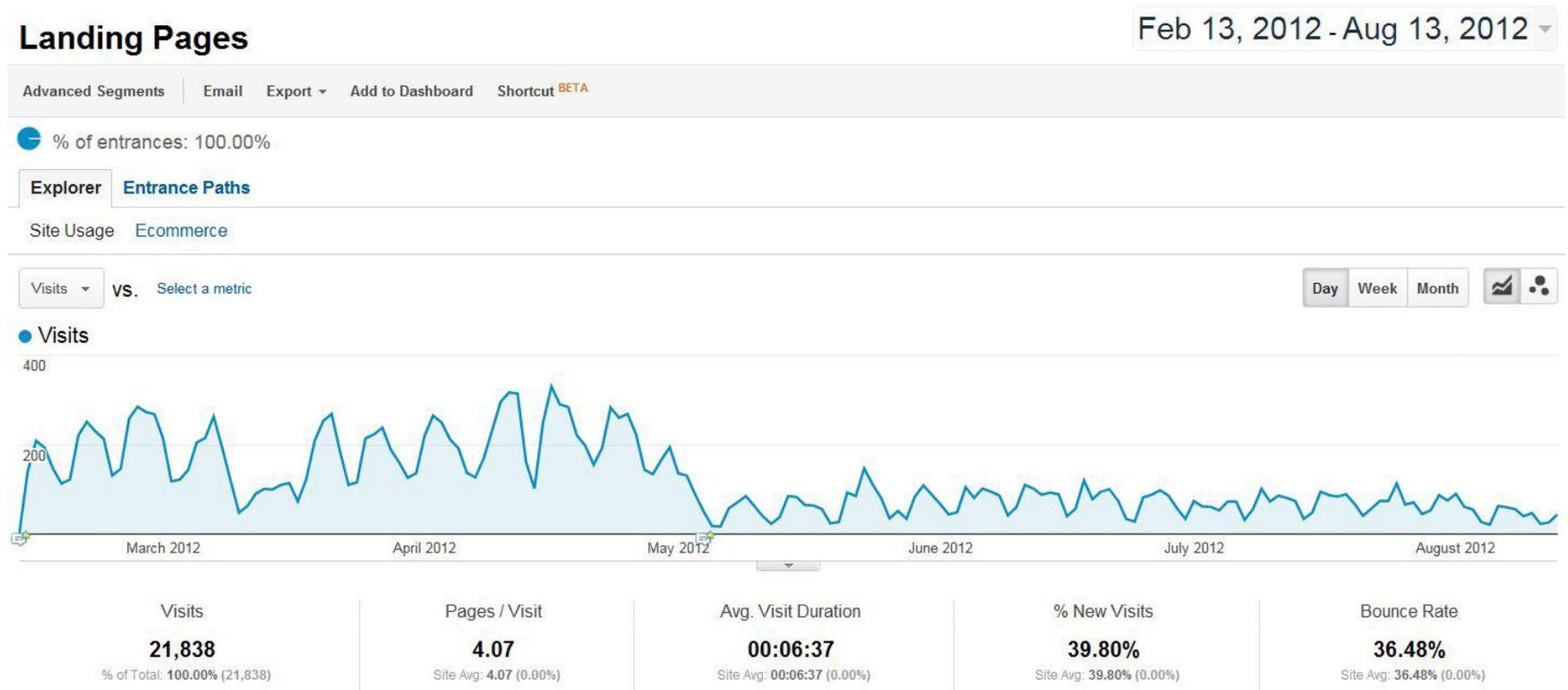


Figure 5: Total landing page visits for MSU Library landing page in six months, 2012.

GA Transaction Log Analysis, continued

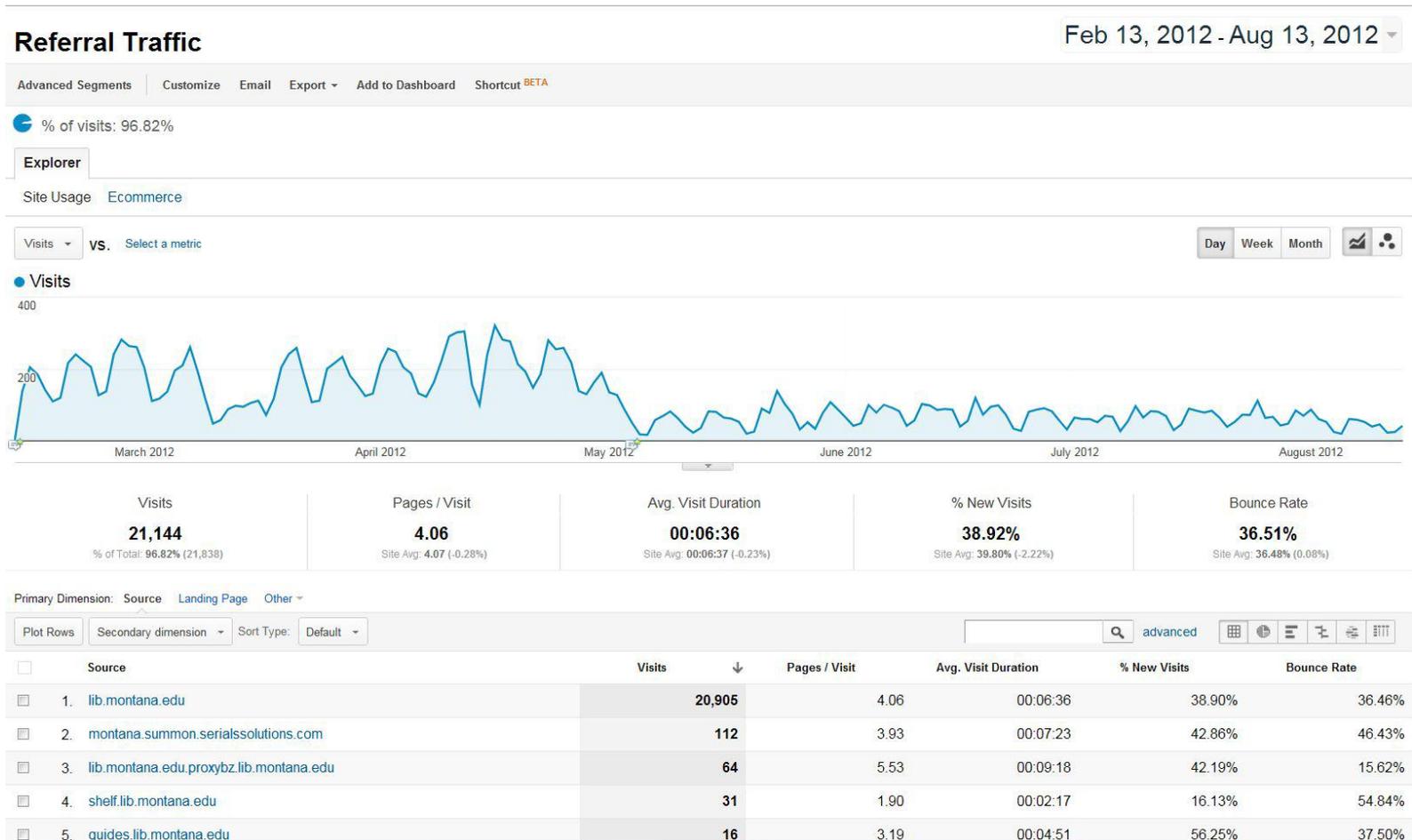


Figure 6: Total referral pages for MSU Library landing page in six months, 2012.

GA Transaction Log Analysis, continued

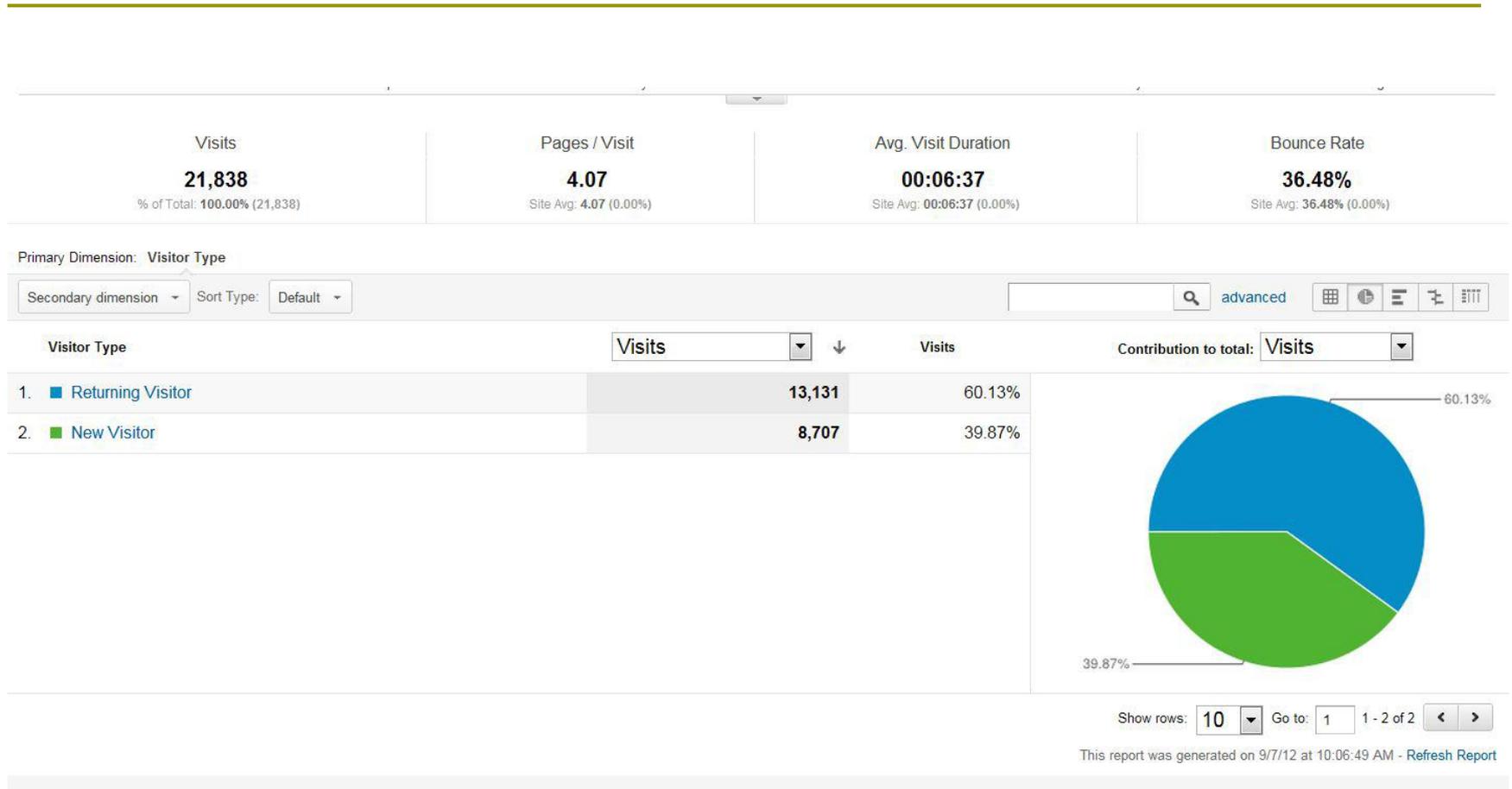


Figure 7: New vs. Returning visitors to Summon

GA Transaction Log Analysis, continued

Pageviews		Unique Pageviews		Avg. Time on Page		Entrances		Bounce Rate		% Exit		Page Value	
88,857		79,997		00:02:09		21,838		36.48%		24.58%		\$0.00	
<small>% of Total: 100.00% (88,857)</small>		<small>% of Total: 100.00% (79,997)</small>		<small>Site Avg: 00:02:09 (0.00%)</small>		<small>% of Total: 100.00% (21,838)</small>		<small>Site Avg: 36.48% (0.00%)</small>		<small>Site Avg: 24.58% (0.00%)</small>		<small>% of Total: 0.00% (\$0.00)</small>	

Primary Dimension: Page Page Title Other												
Plot Rows	Secondary dimension	Sort Type: Default										
<input type="checkbox"/>	Page	Pageviews	Unique Pageviews	Avg. Time on Page	Entrances	Bounce Rate	% Exit	Page Value				
<input type="checkbox"/>	1. /advanced/	3,482	1,957	00:01:01	1,704	19.78%	18.70%	\$0.00				
<input type="checkbox"/>	2. /search?s.q=&x=0&y=0	213	191	00:00:48	164	39.63%	37.09%	\$0.00				
<input type="checkbox"/>	3. /search?s.q=	134	121	00:01:03	35	20.00%	11.94%	\$0.00				
<input type="checkbox"/>	4. /search?s.q=pubmed&x=0&y=0	93	84	00:04:43	77	61.04%	60.22%	\$0.00				
<input type="checkbox"/>	5. /search?s.q=jstor&x=0&y=0	69	62	00:02:43	55	69.09%	68.12%	\$0.00				
<input type="checkbox"/>	6. /	65	56	00:00:58	49	22.45%	20.00%	\$0.00				
<input type="checkbox"/>	7. /search?s.q=CINAHL&x=0&y=0	59	52	00:05:30	50	86.00%	84.75%	\$0.00				
<input type="checkbox"/>	8. /search?s.q=cinahl&x=0&y=0	55	49	00:01:33	45	80.00%	78.18%	\$0.00				
<input type="checkbox"/>	9. /search?s.q=Coelurosauravus	43	42	00:06:23	41	95.12%	95.35%	\$0.00				
<input type="checkbox"/>	10. /search?s.q=hours&x=0&y=0	42	42	00:00:54	40	95.00%	92.86%	\$0.00				
<input type="checkbox"/>	11. /search?s.q=web+of+science&x=0&y=0	39	37	00:07:37	35	85.71%	87.18%	\$0.00				
<input type="checkbox"/>	12. /search	30	22	00:01:13	11	45.45%	33.33%	\$0.00				
<input type="checkbox"/>	13. /search?s.cmd=nextPage()&s.pn=1&s.q=plant+science+master's+thesis+peter+zuck	27	26	00:00:42	25	100.00%	92.59%	\$0.00				
<input type="checkbox"/>	14. /search?s.q=matlab&x=0&y=0	26	23	00:04:13	21	80.95%	80.77%	\$0.00				
<input type="checkbox"/>	15. /search?s.q=pubmed	24	19	00:02:28	5	60.00%	45.83%	\$0.00				
<input type="checkbox"/>	16. /search?s.q=CINAHL	23	21	00:05:02	16	87.50%	86.96%	\$0.00				
<input type="checkbox"/>	17. /search?s.q=library+hours&x=0&y=0	23	22	00:00:53	21	90.48%	91.30%	\$0.00				
<input type="checkbox"/>	18. /search?s.q=mergent&x=0&y=0	21	21	00:01:33	21	85.71%	85.71%	\$0.00				
<input type="checkbox"/>	19. /search?s.q=eric&x=0&y=0	20	19	00:03:11	15	26.67%	45.00%	\$0.00				
<input type="checkbox"/>	20. /search?s.q=cinahl	19	12	00:02:43	8	87.50%	42.11%	\$0.00				

Figure 8: Total pageviews for Summon at MSU Library for six month period

Part 4: Holdings and Indexing

Comparison

- ❑ Comparison of coverage of resources in Summon vs. overall library holdings
- ❑ Implications:
 - If not indexed, should we cancel or consider alternative sources; or, encourage indexing cooperation between data provider and Summon?
 - Do other WSD tools provide better indexing to match our holdings?

Holdings and Indexing, continued

- ❑ Methodology: In August 2012, title-level analysis of holdings against Summon (could have requested at initial implementation, as well, for comparison)
- ❑ Manual comparison of our indexes and databases against Summon's list of full-text coverage

Holdings and Indexing, continued

- ❑ 79,757 serial entries. De-duped by Serials Solutions for 42,464 unique titles. Of those with active ISSNs, only 2,679 were not indexed in Summon and only 709 of those were peer-reviewed sources.
- ❑ Serials Solutions “we can say that we are already in active negotiations with some, if not most, of the content sources on this list.”
- ❑ Summary: 6.3% of our titles aren’t indexed in Summon and only 1.6% of peer-reviewed titles aren’t indexed.

Holdings and Indexing, continued

- We subscribe to 139 databases
- 59 of these are abstracting and indexing databases (so not part of “full-text” coverage) or not appropriate for full-text indexing (such as *ProQuest’s EASI Datasets and Market Planner*)
- 80 databases remaining. 65 are indexed in Summon. 3 of these provided MARC records which are in our library catalog, so are in Summon.
- Summary: 85% of full-text sources in Summon, 15% are not.

Conclusion/Further Study

- ❑ Partnering with other libraries with other WSD tools for comparison of link success
- ❑ Turn off high-offending databases to reduce errors
- ❑ Serve as advocate to get vendors to cooperate, follow standards
- ❑ Consider purpose and placement of WSD
- ❑ Instruction implications
- ❑ Get people to report problems – and follow-up on them
- ❑ User perceptions vs. realities
- ❑ Final question? Did we get what we paid for?