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?
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
Table 1. Failed links to full text

<table>
<thead>
<tr>
<th>Period of study</th>
<th>Percentage failure rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2010</td>
<td>45</td>
</tr>
<tr>
<td>Fall 2011</td>
<td>23</td>
</tr>
<tr>
<td>Summer 2012</td>
<td>27</td>
</tr>
</tbody>
</table>
### Link Analysis, continued

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

<table>
<thead>
<tr>
<th>Description</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Failed Link</td>
<td>158</td>
</tr>
<tr>
<td>Successful Link</td>
<td></td>
</tr>
<tr>
<td>1 click</td>
<td>82</td>
</tr>
<tr>
<td>2 clicks</td>
<td>194</td>
</tr>
<tr>
<td>3 clicks</td>
<td>136</td>
</tr>
<tr>
<td>4 clicks</td>
<td>6</td>
</tr>
<tr>
<td>5 clicks</td>
<td>4</td>
</tr>
<tr>
<td>6 clicks</td>
<td>2</td>
</tr>
<tr>
<td>Non-full text</td>
<td>68</td>
</tr>
<tr>
<td>Total</td>
<td>650</td>
</tr>
</tbody>
</table>
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)
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.

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

<table>
<thead>
<tr>
<th>Query Type</th>
<th>First Year (%)</th>
<th>Second Year (%)</th>
<th>Overall (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>69.8</td>
<td>68.9</td>
<td>69.3</td>
</tr>
<tr>
<td>Known Item</td>
<td>21.7</td>
<td>20.9</td>
<td>21.3</td>
</tr>
<tr>
<td>Database/Journal</td>
<td>5.2</td>
<td>2.8</td>
<td>3.9</td>
</tr>
<tr>
<td>Boolean/Operator</td>
<td>3.3</td>
<td>7.3</td>
<td>5.4</td>
</tr>
</tbody>
</table>
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
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

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?