EVALUATING DISCOVERY:
HOW TOOLS AFFECT UNDERGRADUATES’ RESEARCH PRACTICES

Andrew D. Asher, PhD
Assessment Librarian
Indiana University Bloomington

Recovery & Discovery
Algorithmic Culture

The sorting, classifying, hierarchizing, and curating of people, places, objects, and ideas by machine-based information processing systems (Striphas 2011).

Algorithms as cultural objects

Image credit: http://searchengineland.com/seotable
Embedded value judgments

Non-neutrality

All information technologies favor some content or users over others. One cannot design a neutral system. So, a declaration or description of bias is not an indictment of a system or a firm. A bias is not necessarily bad: it is necessary.

--Siva Vaidhyanathan, *The Googlization of Everything*

---

Discovery Tools Implementation

- 87 students participating:
  - 41 @ Illinois Wesleyan University
  - 46 @ Bucknell University
Discovery Tool Methods

- 5 Test Groups
  - Summon
  - EDS
  - Google Scholar
  - “Conventional” Library Catalog
  - No tool (i.e. Google)

- 4 Research Tasks
  - Find 2 sources per task

- Debriefing Interview
  - Open-ended questions on search practices and evaluation processes

Disclaimer

The goal of this research was not to make purchasing recommendations.
Caveat

Every discovery tool instance is different. Your mileage may vary.

Search Pattern

1. Use something familiar
2. Simple keyword searches
3. First page results
4. Cursory evaluation of sources
Google dominates

Pew Search Engine Use Study

Google is far and away the search engine of choice, preferred by 83% of search users

% of search users who answered the question: Which search engine do you use MOST OFTEN?

- Google: 47% (2004) and 81% (2012)
- Yahoo: 26% (2004) and 4% (2012)
- Other: 17% (2004) and 5% (2012)

Source: The Pew Research Center's Internet & American Life Project Winter 2012 Tracking Survey, January 20-February 25, 2012. N=1,793 adults, age 18 and older, including 833 cell phone interviews. Interview was conducted in English and Spanish.

Google structures expectations

- Single search box
- Simple keyword search

Discovery Tools Project

Simple search was used 82% of the time.

<table>
<thead>
<tr>
<th>Type of Search</th>
<th>Simple</th>
<th>Advanced</th>
<th>Boolean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Scholar</td>
<td>94.5%</td>
<td>4.2%</td>
<td>1.4%</td>
</tr>
<tr>
<td>Summon</td>
<td>79.3%</td>
<td>12.6%</td>
<td>8.1%</td>
</tr>
<tr>
<td>EDS</td>
<td>75.4%</td>
<td>23.1%</td>
<td>1.5%</td>
</tr>
<tr>
<td>Library Catalog/Databases</td>
<td>77.2%</td>
<td>19.1%</td>
<td>3.7%</td>
</tr>
<tr>
<td>No Tool</td>
<td>81.1%</td>
<td>16.3%</td>
<td>2.3%</td>
</tr>
<tr>
<td>Total, All Groups</td>
<td>81.5%</td>
<td>15.1%</td>
<td>3.4%</td>
</tr>
</tbody>
</table>

Source: Discovery Tool Research Results, Illinois Wesleyan University/Bucknell University, 2011
Search Evaluation

- Students iterate search rather than refine.
- Cursory evaluation of sources.
- Eclectic, and sometimes inaccurate, methods of evaluation.
- Assumption that if information is not easily found then it must not exist.

"Apparently you don’t have much on Rock and Roll"
--First Year in French

Discovery Tools Project

92% of the resources utilized were found on the first page of search results.

<table>
<thead>
<tr>
<th>First-page Sources</th>
<th>Percent of sources found on first page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Scholar</td>
<td>83%</td>
</tr>
<tr>
<td>Summon</td>
<td>96%</td>
</tr>
<tr>
<td>EDS</td>
<td>94%</td>
</tr>
<tr>
<td>Library Catalog/Databases</td>
<td>94%</td>
</tr>
<tr>
<td>No Tool Specified</td>
<td>94%</td>
</tr>
</tbody>
</table>

Source: Discovery Tool Research Results, Illinois Wesleyan University/Bucknell University, 2011
Three Biases

De facto outsourcing of evaluation to the search algorithm itself.

Brand Bias  Default Bias  Trust Bias

Brand Bias

Search tool as “surrogate expert”
Trust Bias

“I have no idea [how Google determines search results]. I’m just trusting Google to know what are the good resources.”
--Sophomore in Biology

Default Bias

“I have no idea [how Google determines search results]. I’m just trusting Google to know what are the good resources.”
--Sophomore in Biology
What a tool searches determines what students use:

<table>
<thead>
<tr>
<th>Resource Types</th>
<th>Google Scholar</th>
<th>Summon</th>
<th>EDS</th>
<th>Library Catalog/Databases</th>
<th>No Tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Journal Articles</td>
<td>55.0%</td>
<td>65.0%</td>
<td>73.8%</td>
<td>49.2%</td>
<td>50.3%</td>
</tr>
<tr>
<td>Books</td>
<td>26.5%</td>
<td>13.4%</td>
<td>12.5%</td>
<td>41.3%</td>
<td>15.4%</td>
</tr>
<tr>
<td>Newspapers/Magazines/Trade Journals</td>
<td>2.0%</td>
<td>20.6%</td>
<td>6.3%</td>
<td>3.2%</td>
<td>2.7%</td>
</tr>
<tr>
<td>For-Pay Articles</td>
<td>13.3%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>1.3%</td>
</tr>
<tr>
<td>Websites (including Wikipedia)</td>
<td>0.7%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>21.5%</td>
</tr>
<tr>
<td>Government &amp; Legal Document</td>
<td>2.7%</td>
<td>0.9%</td>
<td>5.0%</td>
<td>2.1%</td>
<td>2.0%</td>
</tr>
<tr>
<td>Other Documents</td>
<td>0.0%</td>
<td>1.0%</td>
<td>2.5%</td>
<td>4.2%</td>
<td>8.7%</td>
</tr>
</tbody>
</table>

Source: Discovery Tool Research Results, Illinois Wesleyan University/Bucknell University, 2011

Library Responses

- Critically assess the effects of search tools
  - Statistical evaluation
  - Modeling of user flows/search terms
Library Responses

- Advocate for Disclosure
  - e.g. NISO Open Discovery Initiative

- Concentrate on teaching algorithmic literacy

- Evaluate implementation of library discovery tools

For more information:

E-mail: asherand@indiana.edu
Twitter: @aasher

Full Results:
Questions?