Use of Traditional News Databases at Penn State University: Trends and Implications

Debora Cheney, The Larry and Ellen Foster Communications Librarian and Head, Library Services to World Campus/PSU Online
The University Libraries
The Pennsylvania State University
Research and Data Support: Alison Ruckert, Foster Librarian Research Assistant

Global Resources Roundtable: Beyond the Fold: Access to News in the Digital Era
June 27, 2013
Newberry Library, Chicago
Three Stories

Aggregator database use
Historical news database use
Impact of discovery systems on news content use
Story 1: Major newspapers; Regional/State newspapers; International newspapers

AGGREGATOR DATABASE
USE
Snapshot: Aggregator Database Use

- Overall usage of Aggregators is declining steadily
- “Major newspaper” use in Aggregators is declining
  - although we have “lost” some data in L/N categories
- *Wall Street Journal/Business* news sources gaining (overlaps with international)
- “Stickiness” of Aggregator titles drive use
- Other trends may be developing: use of wire services and television/news sources
Trend: Database Use by Category, 2006-12

<table>
<thead>
<tr>
<th>Year</th>
<th>Aggregators</th>
<th>Historical</th>
<th>Specialized</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>650066</td>
<td>77639</td>
<td>129</td>
</tr>
<tr>
<td>2007</td>
<td>599467</td>
<td>127687</td>
<td>54</td>
</tr>
<tr>
<td>2008</td>
<td>676187</td>
<td>168007</td>
<td>1330</td>
</tr>
<tr>
<td>2009</td>
<td>521539</td>
<td>153897</td>
<td>4394</td>
</tr>
<tr>
<td>2010</td>
<td>467182</td>
<td>165896</td>
<td>4982</td>
</tr>
<tr>
<td>2011</td>
<td>484778</td>
<td>157295</td>
<td>4823</td>
</tr>
<tr>
<td>2012</td>
<td>459723</td>
<td>210147</td>
<td>8895</td>
</tr>
</tbody>
</table>
**Trend: “Market Share” of each Aggregator Databases**

**Documents Viewed**

<table>
<thead>
<tr>
<th></th>
<th>2010</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factiva</td>
<td>0%</td>
<td>30%</td>
</tr>
<tr>
<td>LexisNexis</td>
<td>43%</td>
<td>40%</td>
</tr>
<tr>
<td>Newsbank</td>
<td>25%</td>
<td>18%</td>
</tr>
<tr>
<td>PQ Nat’l Newsp</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>PQ Ethnic/Black</td>
<td>7%</td>
<td>2%</td>
</tr>
<tr>
<td>WSJ/ABI/Inform</td>
<td>7%</td>
<td>5%</td>
</tr>
</tbody>
</table>
## Trend: % Documents viewed from specific titles (Aggregators only)

<table>
<thead>
<tr>
<th></th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>New York Times</td>
<td>12%</td>
<td>7%</td>
<td>8%</td>
<td>3%</td>
</tr>
<tr>
<td>Wall Street Journal</td>
<td>7%</td>
<td>7%</td>
<td>10%</td>
<td>7%</td>
</tr>
<tr>
<td>Washington Post</td>
<td>4%</td>
<td>2%</td>
<td>2%</td>
<td>1%</td>
</tr>
<tr>
<td>Los Angeles Times</td>
<td>3%</td>
<td>4%</td>
<td>3%</td>
<td>1%</td>
</tr>
<tr>
<td>Chicago Tribune</td>
<td>3%</td>
<td>4%</td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td>USA Today</td>
<td>2%</td>
<td>1%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Other major U.S. cities</td>
<td>15%</td>
<td>6%</td>
<td>1%</td>
<td>1%</td>
</tr>
</tbody>
</table>

**Totals:** 45% 30% 25% 13%
Trend: % Major and PA newspapers Documents Viewed from Aggregators

2009 53%
2010 41%
2011 28%
2012 17%
Stickiness of Aggregator Influences Database Use, 2012

![Bar chart showing the number of documents viewed across different aggregators and news sources in 2012.](image-url)
# International Titles and Other Patterns-2012 (Handout)

## Newsbank
1. Harrisburg Patriot-News
2. Fox News Channel
3. Centre Daily Times
4. CNN
5. Philadelphia Inquirer
6. Pittsburgh Post-Gazette
7. MSNBC
8. Erie Times
9. USA Today
10. Financial Times (London)

## LexisNexis
1. All full text, English News
2. Major World Publications
3. New York Times
4. International News, company, info, etc.
5-6 Misc./Combined categories
7. Globe and Mail (Canada)
8. Washington Post
9. Pittsburgh Post-Gazette
10. Ottawa Citizen

## Factiva
1. Reuters News
2. Agence France-Presse
3. Wall Street Journal
4. Hindu (Abstracts), The
5. Reforma (Spanish Language)
7-9 [Industry pubs]
10. Agence Belga (French Language)
Story 2: Use overall; major papers; collections focus

HISTORICAL NEWS DATABASE
USE
• Use of Historical news databases is significantly lower than use of Aggregators
• Use of historical databases is increasing
• What drives historical database use? Curriculum; Subject-librarian; Research needs likely impact use
• Collection-focus may impact use—PSU focuses on PA titles in databases/collections
• What defines an “historical database”—years or format?
Top Historical Digital Archives, 2012 (Documents viewed)

- PQ New York Times 1851-2010: 74,142
- Readex America’s Historical Newspapers: 67,702
  - Pennsylvania Gazette 10,755
  - Patriot 2,931
  - Philadelphia Inquirer 2,959
- PQ Chicago Tribune 1849-1987: 23,632
- Gale: Burney Collection (17c & 18c): 15,190
- NewspaperArchive: 8,505
- Accessible Archives: 6,097
- Readex African American Newspapers: 5,405
Historical News Content vs. Aggregators, 2009-2012

<table>
<thead>
<tr>
<th>Year</th>
<th>Aggregators</th>
<th>ProQuest Historical Newspapers</th>
<th>Newsbank: American's Historical Newspapers/American Newspaper Archive</th>
<th>Gale</th>
<th>Accessible Archives</th>
<th>NewspaperArchive</th>
<th>Newsbank/CRL: South Asian, Latin American, World Historical</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>79%</td>
<td>10%</td>
<td>6%</td>
<td>5%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>2009</td>
<td>75%</td>
<td>9%</td>
<td>9%</td>
<td>6%</td>
<td>1%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>2010</td>
<td>72%</td>
<td>12%</td>
<td>8%</td>
<td>6%</td>
<td>2%</td>
<td>0%</td>
<td>1%</td>
</tr>
<tr>
<td>2011</td>
<td>74%</td>
<td>12%</td>
<td>8%</td>
<td>5%</td>
<td>1%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>2012</td>
<td>68%</td>
<td>16%</td>
<td>11%</td>
<td>4%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
</tbody>
</table>
Influence what databases are linked to for content

DISCOVERY SYSTEMS
Snapshot: Discovery System Impact

- May affect which database content is linked to and thus documents viewed, searches, etc.
- When database vendor not linked may have to rely on its “stickiness” [and other approaches to attract researchers]
- May increase use of Historical databases when content years overlap with Aggregators
- May impact discovery of historical [if users don’t use search functionality]
Trend: ‘Major Newspaper’ Use Compared
Information, Aggregator Databases 2012
Trend: New York Times

LexisNexis + PQ historical

(documents viewed)

2009: 63,598 + 55,188
2010: 31,810 + 58,656
2011: 39,989 + 57,940
2012: 13,675 + 74,142
Trend: NYT Documents Viewed 2009-2012 (Aggregator Databases)

- **Total**: 118786, 90466, 97929, 87817
- **ProQuest Historical Newspapers: The New York Times (1851-2007)**: 55188, 58656, 57940, 74142
- **ProQuest**: 39823, 11660, 0, 0
- **Newsbank**: 9742, 17312, 17544, 0
- **LexisNexis**: 14033, 2833, 11470, 9634
- **Factiva**: 0, 5, 10975, 4041
Trend: WSJ Documents Viewed, 2010-12 (Aggregator Databases)

<table>
<thead>
<tr>
<th></th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factiva</td>
<td>0</td>
<td>11</td>
<td>1622</td>
<td>4710</td>
</tr>
<tr>
<td>LexisNexis</td>
<td>5</td>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ProQuest</td>
<td>35513</td>
<td>33159</td>
<td>48387</td>
<td>27001</td>
</tr>
<tr>
<td>Newsbank</td>
<td>225</td>
<td>309</td>
<td>170</td>
<td>0</td>
</tr>
</tbody>
</table>
Trend: Searches ≠ Documents Viewed

Rule of Thumb: When newspapers are searched with other non-news content the number of searches increases, but documents viewed decreases.
Other Discovery system impacts

- Don’t distinguish major papers from others
- Don’t distinguish major stories (Page 1)
- Don’t de-dup stories that appear in multiple newspapers
- Too Easy to “remove” newspapers from results
- Don’t search all news databases “equally” or at all—Centre Daily Times
What impacts news content use

IMPLICATIONS [FROM A COLLECTIONS PERSPECTIVE]
Trend Headline: License a Lot, Use a Little, Use declining

- Aggregator database use is declining; “historical” increasing
- Major title use is strong, but declining
- Database use is driven by “stickiness”
- Discovery systems are probably impacting database and title use
Libraries Role as a News Content Provider Eroding

- Users don’t associate libraries with news content
- Users have to work too hard to find news content
- News doesn’t mean newspapers anymore
- We don’t have the news content they want/need
- Faculty (not librarians) influence where and how their students search for news content
Libraries and News Content

Possibly Too Tied to the Concept of the “newspaper”
Questions, Comments, Suggestions?

Debora Cheney, The Larry and Ellen Foster Communications Librarian and Head, Library Services to World Campus/PSU Online

The University Libraries
The Pennsylvania State University
Pennsylvania
dcheney@psu.edu