What’s Asked Where?

Modelling a Large Reference Interaction Dataset

Andrew Asher, PhD
Assessment Librarian
@aasher
asherand@Indiana.edu
IUB Libraries Reference Data

- 457,794 Logged Reference Questions, 2006-2013
- 16,337 E-mail Transcripts, 2007-2013
  - 31,109 Chat Transcripts, 2006-2012

- Data from the main library, Herman B Wells (B-Wells) & 16 branch libraries.
Outline

• Demonstrate strategies for approaching a large dataset.
• Present preliminary results from Indiana University

• A quantitative approach to qualitative insights
31%

Decrease in reference questions since 2006

-44%
Main library
-12%
All other libraries
-2723
Questions per year
IUB Libraries
Reference Transactions
Average Yearly Gain/Loss (in absolute numbers)
## Overall Hourly Reference Demand 2006-2013

<table>
<thead>
<tr>
<th>Hour</th>
<th>Sunday</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>40</td>
<td>117</td>
<td>201</td>
<td>156</td>
<td>128</td>
<td>109</td>
<td>65</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>15</td>
<td>32</td>
<td>7</td>
<td>18</td>
<td>25</td>
<td>26</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
<td>7</td>
<td>9</td>
<td>6</td>
<td>9</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>10</td>
<td>10</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>37</td>
<td>29</td>
<td>27</td>
<td>122</td>
<td>34</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>28</td>
<td>700</td>
<td>913</td>
<td>994</td>
<td>855</td>
<td>783</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>2,367</td>
<td>2,583</td>
<td>2,453</td>
<td>2,414</td>
<td>2,107</td>
<td>32</td>
</tr>
<tr>
<td>9</td>
<td>17</td>
<td>4,746</td>
<td>5,600</td>
<td>5,045</td>
<td>4,620</td>
<td>3,857</td>
<td>168</td>
</tr>
<tr>
<td>10</td>
<td>21</td>
<td>6,932</td>
<td>7,080</td>
<td>6,332</td>
<td>6,206</td>
<td>5,545</td>
<td>1,446</td>
</tr>
<tr>
<td>11</td>
<td>1,107</td>
<td>8,422</td>
<td>9,064</td>
<td>8,325</td>
<td>7,618</td>
<td>6,775</td>
<td>1,757</td>
</tr>
<tr>
<td>12</td>
<td>1,888</td>
<td>8,763</td>
<td>8,623</td>
<td>8,426</td>
<td>8,100</td>
<td>6,579</td>
<td>2,327</td>
</tr>
<tr>
<td>13</td>
<td>3,450</td>
<td>9,367</td>
<td>8,622</td>
<td>7,923</td>
<td>7,658</td>
<td>7,730</td>
<td>3,666</td>
</tr>
<tr>
<td>14</td>
<td>3,967</td>
<td>9,334</td>
<td>9,887</td>
<td>9,000</td>
<td>8,178</td>
<td>6,976</td>
<td>3,109</td>
</tr>
<tr>
<td>15</td>
<td>4,299</td>
<td>9,474</td>
<td>9,861</td>
<td>9,394</td>
<td>7,986</td>
<td>6,746</td>
<td>3,339</td>
</tr>
<tr>
<td>16</td>
<td>4,056</td>
<td>8,876</td>
<td>8,741</td>
<td>8,982</td>
<td>7,132</td>
<td>7,366</td>
<td>3,869</td>
</tr>
<tr>
<td>17</td>
<td>3,516</td>
<td>7,340</td>
<td>7,006</td>
<td>6,693</td>
<td>6,046</td>
<td>3,586</td>
<td>1,856</td>
</tr>
<tr>
<td>18</td>
<td>2,941</td>
<td>4,566</td>
<td>4,105</td>
<td>4,186</td>
<td>3,375</td>
<td>1,351</td>
<td>1,223</td>
</tr>
<tr>
<td>19</td>
<td>2,013</td>
<td>3,797</td>
<td>3,773</td>
<td>3,791</td>
<td>2,530</td>
<td>1,011</td>
<td>948</td>
</tr>
<tr>
<td>20</td>
<td>2,638</td>
<td>3,934</td>
<td>3,016</td>
<td>3,355</td>
<td>2,574</td>
<td>1,276</td>
<td>1,677</td>
</tr>
<tr>
<td>21</td>
<td>2,256</td>
<td>3,366</td>
<td>3,178</td>
<td>2,911</td>
<td>2,256</td>
<td>98</td>
<td>128</td>
</tr>
<tr>
<td>22</td>
<td>1,591</td>
<td>2,146</td>
<td>1,788</td>
<td>1,868</td>
<td>1,269</td>
<td>105</td>
<td>123</td>
</tr>
<tr>
<td>23</td>
<td>1,086</td>
<td>1,387</td>
<td>1,448</td>
<td>1,333</td>
<td>1,053</td>
<td>123</td>
<td>77</td>
</tr>
</tbody>
</table>
Questions involving books accounted for 26.2% of reference questions.

Articles and journals accounted for 11.4%
Reference Effort Assessment Data (READ)

A six-point scale evaluating the effort required to answer a reference question.

1 = Simple directional question
6 = In-depth research consultation with faculty or graduate student

For more information see http://readscale.org/
Mean = 2.3
St. Dev. = .817
N = 901
ICC = .905
(High Agreement)
ANOVA: $F = 11.328, p < .01$
\[ \eta^2 = .048 \]

ANOVA: $F = 3.115, p < .01$
\[ \eta^2 = .026 \]

ANOVA: $F = 4.106, p < .01$
\[ \eta^2 = .057 \]
ANOVA: $F = 9.698, p < .01$
$\eta^2 = .044$

**Wells vs. All Other Libraries**
ANOVA: $F=32.156, p > .336$ (not significant)

**Archives vs. All Other Libraries**
ANOVA: $F=32.156, p < .01$
$\eta^2 = .035$
Topic Modeling

• Probabilistic approach to inferring themes across a corpus of documents by the frequency of co-occurring words.
List of Topics

1. book title request copy iucat books number author loan english item checked history interlibrary ill published location edition available notes
2. site search login education source direct article study live research link ebscohost technology social school true lib aspx use national
3. page search pdf not available we id it lib cpi ok so possible section then pages web html let
4. know help looking need like paper time research let finding sources published articles think dissertation way best doing appreciated look
5. access article journal online trying link search articles page database journals using able click website campus iucat issue password tried
6. student like books wondering check school class time graduate professor studies know form course music returned working possible id need
7. message phone th original question librarian address ask resources fax information mary cc monday wrote wednesday electronic thursday services june
8. new information year best research list proquest business help american state international need looking like magazine time world york news
9. question information use students mail ask librarian account access service help send services faculty databases address resources research staff contact
10. xml pk bit item ed bie bitu alt ao th oo dw dr cw red ds ok cc cc dj

Example Stopwords: i'm, it's, iu, bloomington
Hi, cw, dmk, iupui, hello, Indiana, iub, thank, thanks, edu Libref, herman, wells, thanks, http, https
Questions

asherand@Indiana.edu
@aasher