FUSE LABS

Future Social Experiences Labs
Lili Cheng
July 2012
Realtime Social Data

August 2009: Twitter data: Public status messages
- 12 million updates a day (20x our estimates)
- Massive growth every month

October 2010: Twitter data + Facebook public data
- 500 million updates a day
- 15,000 storage requests a second

2 guys + Windows Azure + Bing
Link Expansion and Analysis

- Short URL expansion
- Content type
- Language
- Thumbnail
- Images
- Adult score
- Spam score
- (API aware)
Social Search

- Bing Now
- Twitter
- Facebook
- Bing Answers
- Indexing urls

....
...thanks FAST!

- Language detection
- Parts of speech
- Nouns, verbs
- Questions
- Entity extraction
- People
- Places
- Topics
- Custom filters
Social Apps and Components

Profile Service
- Multiple Logins
- Single Profile
- Customizable UI
- Simple REST access

Recommendation Service
- Javascript based search library
- Normalizes all sources to a common schema
- Download only the sources you use
- Easily extended to support new sources
- Support Rx.js async observable pattern
- Sources include RSS, OData, OpenGraph

jSearch
- Javascript based search library
- Normalizes all sources to a common schema
- Download only the sources you use
- Easily extended to support new sources
- Support Rx.js async observable pattern
- Sources include RSS, OData, OpenGraph

URL Analysis
- Classification: adult, article, audio, bookmark, comment, event, file, folder, group, language, list, note, person, photo, album, playlist, product, review, service, status, video
- Extract meta-data from OpenGraph and MicroFormats
- Handles URL shorteners like bit.ly
- Generates web page thumbnails
- Extract media
- Many Sources: Facebook, YouTube, Bing, Google Maps, etc.

Distributed Search
- REST API Store & Retrieve JSON
- Optional Schema with validation
- Full-text and relational search

eSearch: Embedded Search & Storage Library
- Robust, Plugable Item Storage
- 34 simple data types and vectors
- Relational: <, <=, =, >, >=
- Text: contains, near, phrase
- Ranking: explicit, hits and IR
- Multi-level sorting
Authoring. Montage.
So.cl

*Use search to express & share your ideas*

- So.cl
  - Integrates Search & Social
  - Data reuse
    - Bing API’s
    - So.cl Public data & API’s
- About 300,000 users
Springtime in the Southwest

Dan Marshall: I love the springtime in the Southwest US. The blooming flowers speckle and turn it into an amazing place. 2 min ago

Dan Marshall: You always post the most beautiful images. I love the southwest. 2 min ago

Tom Jones: I love that area too. 2 min ago

27 comments

Crazy Cats

Dan Marshall: Searched. 2 min ago

Dan Marshall: You always post the most beautiful images. I love the southwest. 2 min ago

Tom Jones: I love that area too. 2 min ago

27 comments

Carrots

Dan Marshall: Nothing like the orange bits. 2 min ago

Dan Marshall: You always post the most beautiful images. I love the southwest. 2 min ago

Tom Jones: I love that area too. 2 min ago
Factors about equally predict if user comes back

### Component Matrix

<table>
<thead>
<tr>
<th>Type</th>
<th>Creators</th>
<th>Socialites</th>
<th>Browsers</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Variance:</td>
<td>32%</td>
<td>12%</td>
<td>9%</td>
</tr>
<tr>
<td>Created post</td>
<td>.86</td>
<td>.17</td>
<td>.10</td>
</tr>
<tr>
<td>Invited</td>
<td>.01</td>
<td>-.16</td>
<td>.63</td>
</tr>
<tr>
<td>Followed</td>
<td>-.03</td>
<td>.10</td>
<td>.37</td>
</tr>
<tr>
<td>Added item to post</td>
<td>.83</td>
<td>.08</td>
<td>-.06</td>
</tr>
<tr>
<td>Searched</td>
<td>.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commented</td>
<td>.36</td>
<td>.64</td>
<td>.09</td>
</tr>
<tr>
<td>Liked post</td>
<td>.15</td>
<td></td>
<td>.58</td>
</tr>
<tr>
<td>Liked comment</td>
<td>.13</td>
<td>.50</td>
<td>.06</td>
</tr>
<tr>
<td>Messaged</td>
<td>-.09</td>
<td>.50</td>
<td>-.08</td>
</tr>
<tr>
<td>Viewed person</td>
<td>.22</td>
<td>.47</td>
<td>.48</td>
</tr>
<tr>
<td>Navigated to All</td>
<td>.51</td>
<td>.37</td>
<td>.53</td>
</tr>
<tr>
<td>Joined party</td>
<td>.17</td>
<td>.09</td>
<td>.68</td>
</tr>
</tbody>
</table>

### Regression Coefficients

<table>
<thead>
<tr>
<th>Type</th>
<th>Beta</th>
<th>t</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creating</td>
<td>.14</td>
<td>5.28</td>
<td>.000</td>
</tr>
<tr>
<td>Socializing</td>
<td>.07</td>
<td>2.61</td>
<td>.000</td>
</tr>
<tr>
<td>Browsing</td>
<td>.19</td>
<td>7.20</td>
<td>.000</td>
</tr>
</tbody>
</table>

Browsing stronger predictor of overall activity level
Usage: Return Visits

Social responses inspire people to return to site, especially if occurring during first session

- Received Response During First Session
  - No: Proportion Came Back
  - Yes: Proportion Came Back

- Received Response Between First Session and Second Session
  - No: Proportion Came Back
  - Yes: Proportion Came Back

N = 1268  N = 111  N = 1083  N = 293

Social responses to user: following, commenting, liking post, liking comment, riffing
Lessons Learned

Social networking feed:
- Appropriate “flow of data” primary part of social user experience
- Critical mass of activity & “good users” define site norms and personality
- Discovery of new people and topics a person cares about key

Social across “boundaries”:
- Across identities:
- Across time: interaction history creates rich profile over time
- Reuse of public data
Faculty Summit Day 3!
iSchool Social Media Research
Student Social Media Advisors Program
http://fuse.microsoft.com/ssmap

Contacts:
Shelly Farnham  shellyf@microsoft.com
Lee Dirks  ldirks@microsoft.com
Lili Cheng  lilich@microsoft.com
lilich@Microsoft.com
@lilich

http://fuse.microsoft.com
http://so.cl