Social Collaboration Data

Fuzzy Intersects In the Open World...

HP Labs: Sujata Millick, Raj Rajagopalan
Sujata.millick@hp.com
Raj.rajagopalan@hp.com
3rd Billion

<table>
<thead>
<tr>
<th>Connectivity</th>
<th>Pop.</th>
<th>Pentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>1 B</td>
<td>11%</td>
</tr>
<tr>
<td>Asia</td>
<td>3.8B</td>
<td>22%</td>
</tr>
<tr>
<td>Europe</td>
<td>0.8B</td>
<td>59%</td>
</tr>
<tr>
<td>Middle East</td>
<td>0.2B</td>
<td>30%</td>
</tr>
<tr>
<td>North America</td>
<td>0.3B</td>
<td>78%</td>
</tr>
<tr>
<td>Latin America/Caribbean</td>
<td>0.6B</td>
<td>35%</td>
</tr>
<tr>
<td>Oceania/Australia</td>
<td>0.04B</td>
<td>61%</td>
</tr>
</tbody>
</table>

*1.1B, 65%  4.8B, 19.7%  0.8B, 6%

Modes of Social Contact
- Literacy
- Numeracy
- Pictography

Betterment Uses
- Economic
- Entertainment
- Relationships

Language
- 3000+ languages in Africa (12 100-language clusters spoken by 75%)
- 57 official languages – Asia (Central, SouthEast), Middle East
- China – 8 language families, 292 living languages
- India – 1600 mother tongues, 415 recognized.

* July 2010 stat
Changing Nature of Data...

Structured Data
- Meta data
- Sensor data
- Sorted, Organized
- Analyzed using traditional statistical techniques
- Created by formal organizations

Unstructured Data
- Text
- Audio
- Video
- Analysis requires domain, context, sentiment info
- Created “primarily” by individual actors

Data Growth: 2009-2020

2009
0.8 ZBs

*1 Zettabyte = 1 trillion gigabytes

2020
35 ZB*

What’s the value of “he knew before the world knew”?
Turkey earthquake: Twitter plea for help gets 17,000 responses

Journalist Erhan Çelik asks his Twitter followers to offer accommodation to those made homeless by the earthquake

Constanze Letsch in Istanbul and Peter Walker
guardian.co.uk, Monday 24 October 2011 06.57 EDT
Article history

Turkish people sit on a street in Ercis after the 7.2 magnitude earthquake on Sunday. Photograph: Mustafa Ozer/AFP/Getty Images
Data $\rightarrow$ Information

Traditional
- Statistical Analytics
- Correlations
- Select Population Surveys
- Experiments

New...
- Blogging
- Walls
- Tagging
- Tweeting
- Crowdsourcing
- Mashups
- Retweets

ACTION

- Tunisia, Egypt, and other
  - Information to Action predominated
  - Bypassed the “data” stage
  - Data analytics after the event
- Japan Tsunami...
  - Govt. transparency after Information
The Long Tail of Social Media?
Social Web and Peace/Conflict

www.Peace.Facebook.com

Stanford University: Persuasive Technologies Institute (Margarita Quihuis, Mark Nelson)
Social Web and Disaster Response

Haiti Earthquake Relief Project: 4626
Processed 100,000 SMS messages
5000 in 1 hr (peak volume)
Average response time: 2 minutes

Modes: SMS text, video, photo, online report
Social Web and Security
Behavior → Technology

Behaviors

Base of the Pyramid
North America
Europe

Middle of the Pyramid
Emerging

Technology → Behavior

China?
India?
Middle East?
Rest of Asia?
Africa?

Crossing the Chasm, Geoffrey Moore
Hyper-Assumptions

• The 3rd Billion will use the web differently
• The Social Web will be the dominant collaboration and content creation mechanism
• Technology will be shaped by social web-interactions, which will be driven by local cultural mores.
Global IT Revenue: $1.3 Trillion (2009)

U.S. Companies IT Revenue: $800 Billion (2009)

Silicon Valley Companies IT Revenue: $500B (2009)

U.S. IT Companies R&D spending: $30 Billion (2009)

Global IT Security Spend: $2.7 Billion (2010 est.)

R&D/Global Revenue: 2.3%
Ongoing work...

February 2012 Workshop on Social Media and (Security) or (Knowledge Democracy...) Defense, HP Labs, Stanford, Security-Innovation.org, NPS...

Fleshing out the 3rd Billion – data effort (Nov-Dec 2011) Meetup @ Google – Nov 15, 2011

Grassroots engagement – Open World
Possible NPS engagement?

NPS – boundary spanning (Silicon Valley and DC)
Students, researchers, co-collaborators
Class projects, lectures

Follow-up comments, ideas, questions
Sujata Millick, Raj Rajagopalan
Sujata.millick@hp.com
Raj.rajagopalan@hp.com

THANK YOU!
SO WHAT ??

WHO: 3rd Billion
WHEN: 2+ years
WHAT: Content, collaborations
WHERE: Social web
WHY: Culture, language, different needs
HOW: Devices, Apps
# WORLD INTERNET USAGE AND POPULATION STATISTICS

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>1,013,779,050</td>
<td>4,514,400</td>
<td>110,931,700</td>
<td>10.9 %</td>
<td>2,357.3 %</td>
<td>5.6 %</td>
</tr>
<tr>
<td>Asia</td>
<td>3,834,792,852</td>
<td>114,304,000</td>
<td>825,094,396</td>
<td>21.5 %</td>
<td>621.8 %</td>
<td>42.0 %</td>
</tr>
<tr>
<td>Europe</td>
<td>813,319,511</td>
<td>105,096,093</td>
<td>475,069,448</td>
<td>58.4 %</td>
<td>352.0 %</td>
<td>24.2 %</td>
</tr>
<tr>
<td>Middle East</td>
<td>212,336,924</td>
<td>3,284,800</td>
<td>63,240,946</td>
<td>29.8 %</td>
<td>1,825.3 %</td>
<td>3.2 %</td>
</tr>
<tr>
<td>North America</td>
<td>344,124,450</td>
<td>108,096,800</td>
<td>266,224,500</td>
<td>77.4 %</td>
<td>146.3 %</td>
<td>13.5 %</td>
</tr>
<tr>
<td>Latin America/Caribbean</td>
<td>592,556,972</td>
<td>18,068,919</td>
<td>204,689,836</td>
<td>34.5 %</td>
<td>1,032.8 %</td>
<td>10.4 %</td>
</tr>
<tr>
<td>Oceania / Australia</td>
<td>34,700,201</td>
<td>7,620,480</td>
<td>21,263,990</td>
<td>61.3 %</td>
<td>179.0 %</td>
<td>1.1 %</td>
</tr>
<tr>
<td><strong>WORLD TOTAL</strong></td>
<td><strong>6,845,609,960</strong></td>
<td><strong>360,985,492</strong></td>
<td><strong>1,966,514,816</strong></td>
<td><strong>28.7 %</strong></td>
<td><strong>444.8 %</strong></td>
<td><strong>100.0 %</strong></td>
</tr>
</tbody>
</table>

**NOTES:** (1) Internet Usage and World Population Statistics are for June 30, 2010.