

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

Connectivity	Pop.	Petration
Africa	1 B	11%
Asia	3.8B	22%
Europe	0.8B	59%
Middle East	0.2B	30%
North America	0.3B	78%
Latin America/Caribbean	0.6B	35%
Oceania/Australia	0.04B	61%

Modes of Social Contact

Literacy
Numeracy
Pictography

Betterment Uses

Economic
Entertainment
Relationships

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



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...

"The Digital Universe Decade: Are You Ready?" (IDC – Sponsored by EMC, May 2010).

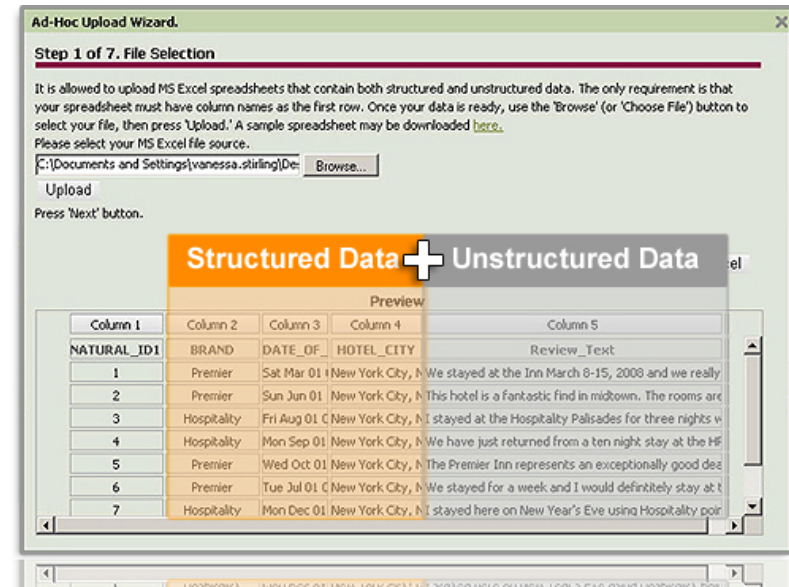
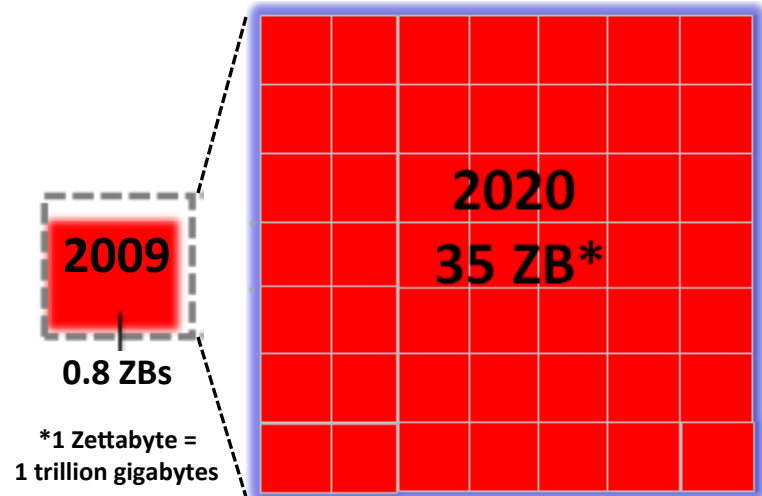
Structured Data

- 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



twitter  Search Home Profile Mes



Sohaib Athar

@ReallyVirtual Abbottabad Lahore Pakistan
An IT consultant taking a break from the rat-race by hiding in the mountains with his laptops.
<http://www.reallyvirtual.com>

[+ Follow](#)  

Timeline Favorites Following ▾ Followers ▾ Lists ▾



ReallyVirtual Sohaib Athar
Here's a picture of OBL's hideout in Abbottabad, as shared by a friend @Rahat <http://yfrog.com/h7w4izmj>
7 minutes ago



ReallyVirtual Sohaib Athar
I apologize for reporting the operation 'unwittingly/unknowingly' - had I known about it, I would have tweeted about it 'wittingly' I swear.
21 minutes ago



ReallyVirtual Sohaib Athar
I just woke up and need to take care of some business before I (re)start blogging - please don't call me just yet. Thank you!
1 hour ago



ReallyVirtual Sohaib Athar

Abbottabad

Twitter User: #ReallyVirtual

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

[Tweet](#) 466

[Recommend](#) 373

[reddit this](#)

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



[A](#) larger | smaller

World news

Turkey · Natural disasters and extreme weather · Middle East · Europe

Technology

Twitter · Internet

Media

Blogging

More news

More on this story



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

Twitter Hashtag:
#EvimEvindirVan

“ my home is your
home, Van”

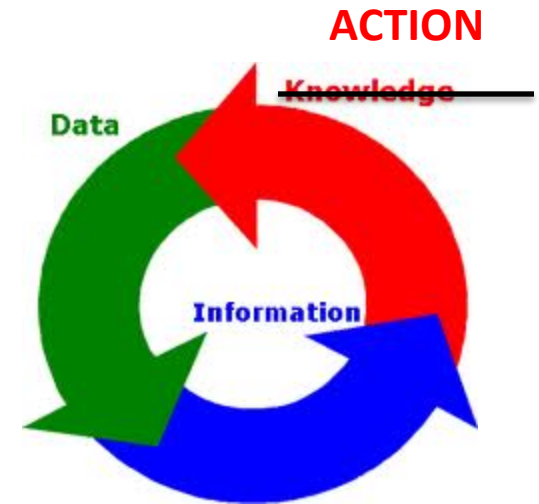
October 23, 24 2011



Data → Information

Traditional

- Statistical Analytics
- Correlations
- Select Population Surveys
- Experiments



New...

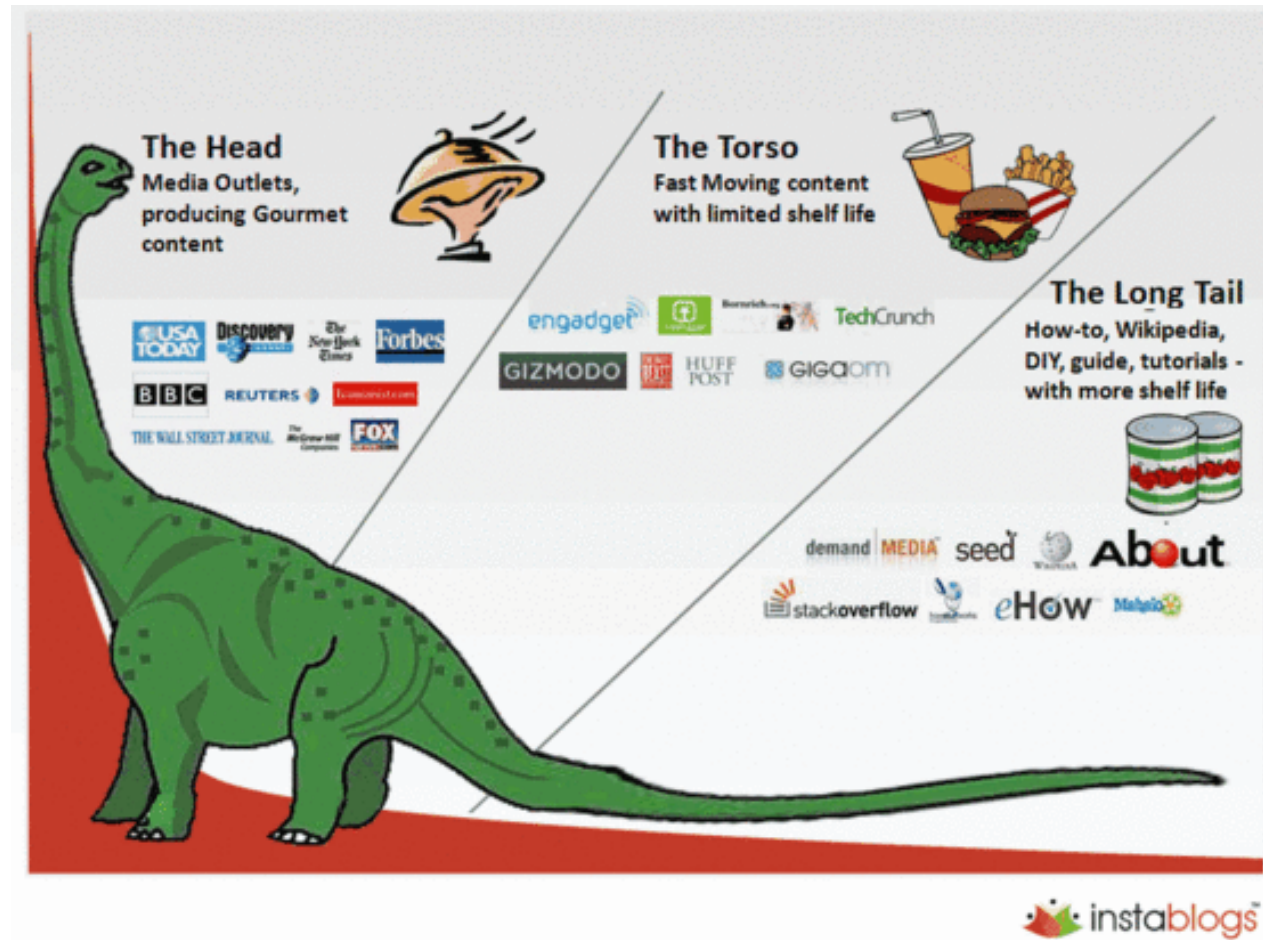
- Blogging
- Walls
- Tagging
- Tweeting
- Crowdsourcing
- Mashups
- Retweets

Data ↔ Information

- 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

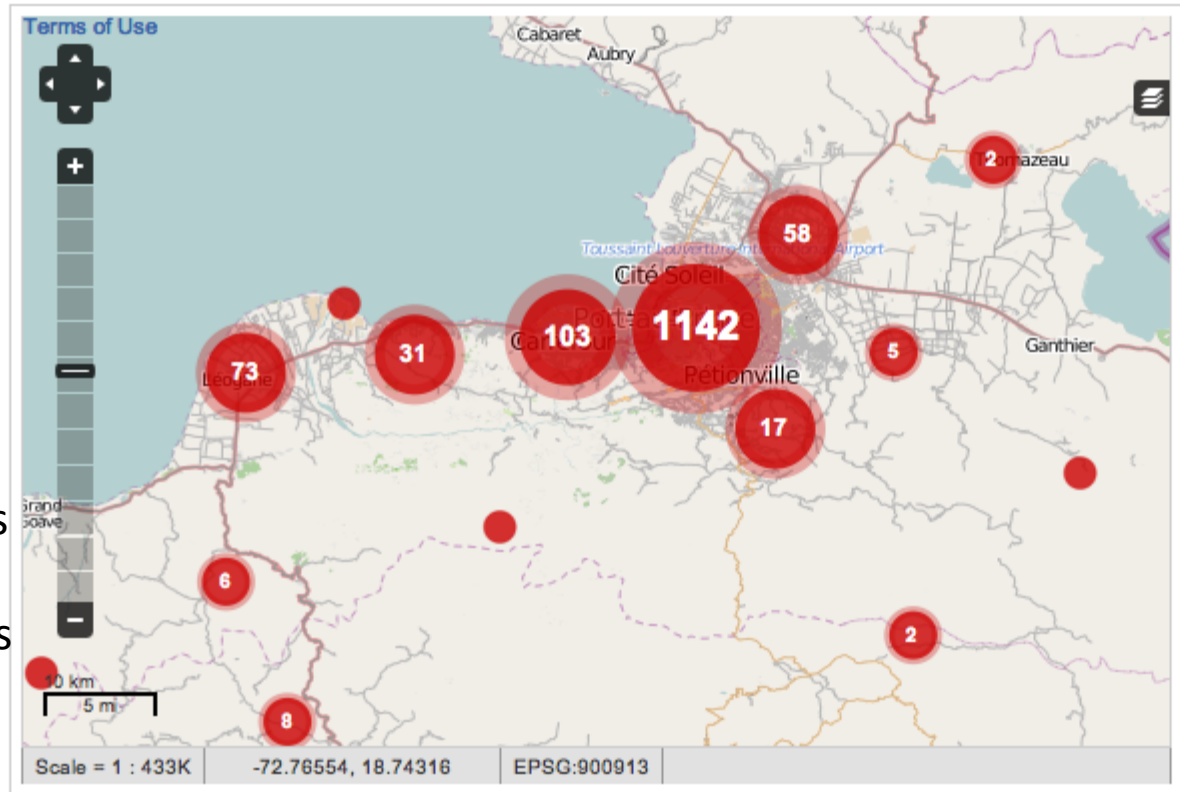
Haiti Earthquake Relief

Project: 4626

Processed 100,000 SMS messages

5000 in 1 hr (peak volume)

Average response time: 2 minutes

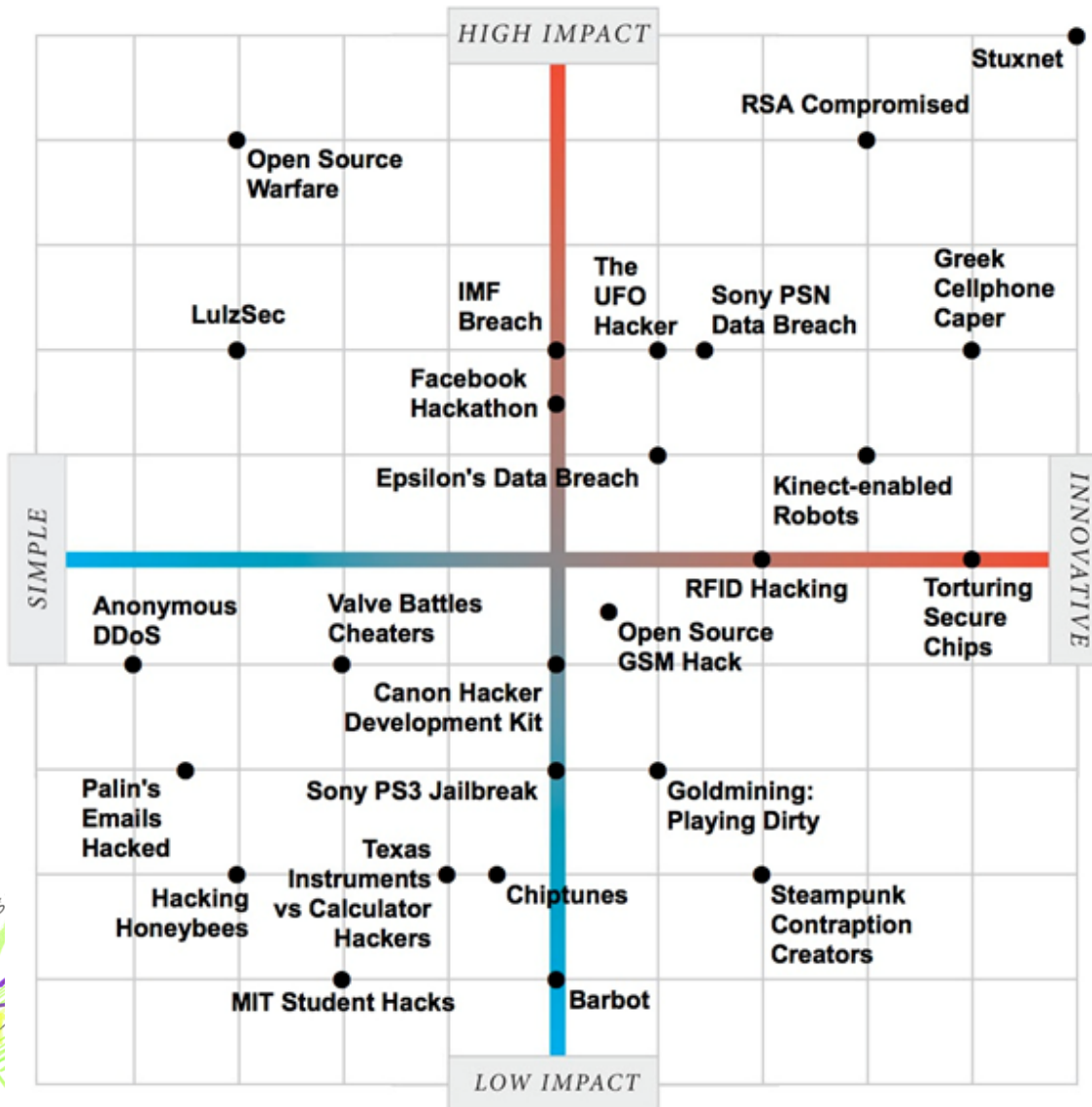


CrowdFlower



Modes: SMS text, video, photo, online report

Social Web and Security

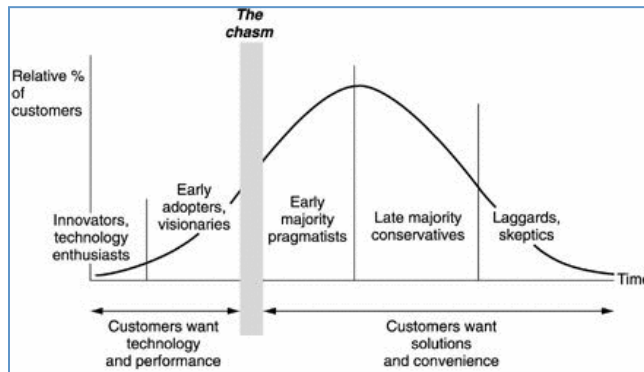


Anonymous

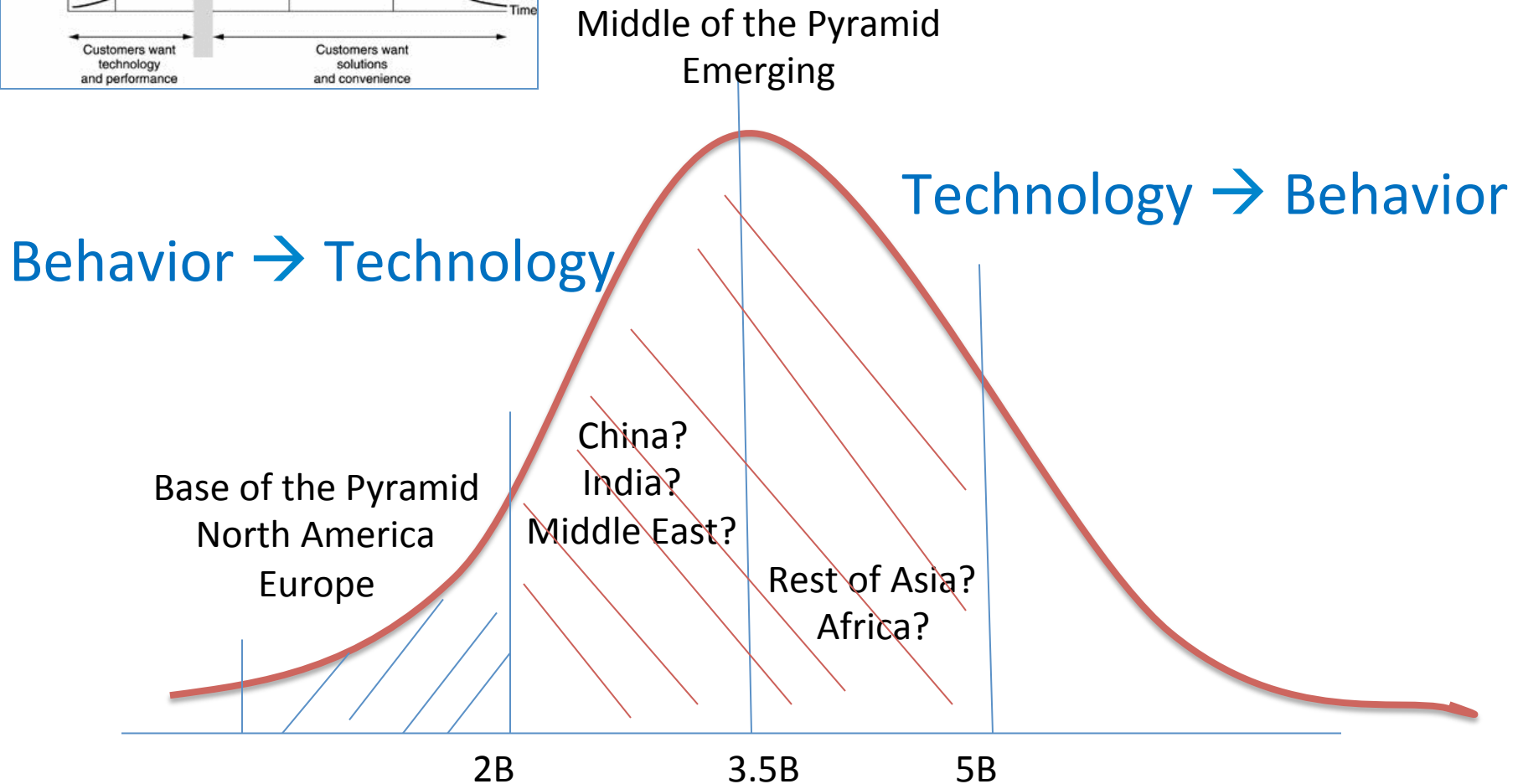


A flag conveying symbolism associated with Anonymous. The imagery of the "suit without a head" represents leaderless organization and anonymity.^[1]





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.



2 Billion people online

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)

R&D/Global Revenue: 2.3%

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

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

World Regions	Population (2010 Est.)	Internet Users Dec. 31, 2000	Internet Users Latest Data	Penetration (% Population)	Growth 2000-2010	Users % of Table
Africa	1,013,779,050	4,514,400	110,931,700	10.9 %	2,357.3 %	5.6 %
Asia	3,834,792,852	114,304,000	825,094,396	21.5 %	621.8 %	42.0 %
Europe	813,319,511	105,096,093	475,069,448	58.4 %	352.0 %	24.2 %
Middle East	212,336,924	3,284,800	63,240,946	29.8 %	1,825.3 %	3.2 %
North America	344,124,450	108,096,800	266,224,500	77.4 %	146.3 %	13.5 %
Latin America/Caribbean	592,556,972	18,068,919	204,689,836	34.5 %	1,032.8 %	10.4 %
Oceania / Australia	34,700,201	7,620,480	21,263,990	61.3 %	179.0 %	1.1 %
WORLD TOTAL	6,845,609,960	360,985,492	1,966,514,816	28.7 %	444.8 %	100.0 %

← Asia

← Europe

← N. America

79.7%

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