The New Paradigm Shift: The Emergence of Social Computing

Irwin King

Department of Computer Science and Engineering
The Chinese University of Hong Kong
http://www.cse.cuhk.edu.hk/~king

Social Networking

HOW TO USE WEB 2.0 IN THE ENTERPRISE

PART 1:
COMMUNICATE WITH YOUR EMPLOYEES

Billionaires' Shuffle

2008

Facebook in 2004.02

2008

at **23** and $**1.5** billion later...

2009

Alexa as of Nov. 2008	China	USA	Japan	India	Brazil	Global
1	Baidu	Google	Yahoo.jp	Google.in	Google	Google
2	QQ	Yahoo	FC2	Google	Orkut.br	Yahoo
3	Sina	Facebook	Google.jp	Yahoo	Windows Live	YouTube
4	Google.cn	YouTube	YouTube	Orkut.in	Universo Online	Facebook
5	Taobao	Myspace	Rakuten	YouTube	YouTube	Windows Live
6	163	MSN	Livedoor	Blogger	Globo	MSN
7	Google	Windows Live	Ameblo.jp	Rediff	MSN	Wikipedia
8	Sohu	Wikipedia	mixi	Facebook	Google	Blogger
9	Youku	Craigslist	Wikipedia	Wikipedia	Yahoo	Baidu
10	Yahoo	EBay	Google	Windows Live	Terra	Myspace

China's Great Firewall

HOME PAGE | TODAY'S PAPER | VIDEO | MOST POPULAR | TIMES TOPICS

The New York Times
Friday, June 19, 2009

News

Search All NYTimes.com Go

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION | ARTS | STYLE | TRAVEL | JOBS | REAL ESTATE | AUTOS

The Lede

The New York Times News Blog

June 2, 2009, 7:05 PM

China's Great Firewall Blocks Twitter

By ROBERT MACKEY

Catherine Henriette/Agence France-Presse — Getty Images

Search This Blog Search

Previous Post: **Bloggers Ponder Last Message From Missing Jet's Computer**

Next Post: **Punditry From Bin Laden and Zawahiri on Obama's Trip to the Middle East**

Recent Posts

June 18 (38 comments)
Latest Updates on Iran's Disputed Election
To supplement reporting from New York Times correspondents inside Iran on Thursday, The Lede will continue to track the aftermath of Iran's disputed presidential election online.

June 17 (129 comments)
Wednesday: Latest Updates on Iran's Disputed Election
On Wednesday, The Lede will continue to track the aftermath of Iran's disputed presidential election online, to supplement reporting from New York Times correspondents inside Iran.

June 16 (198 comments)
Tuesday: Latest Updates on Iran's Disputed Election
To supplement reporting from New York Times correspondents inside Iran, The Lede

Twitter in Iran's Revolution

twitter

Login Join Twitter!

Rallying Iran: Time Tempers a Challenger Forged in Revolution
http://bit.ly/epfBT

5:45 PM Jun 17th from web

nytimes
The New York Times

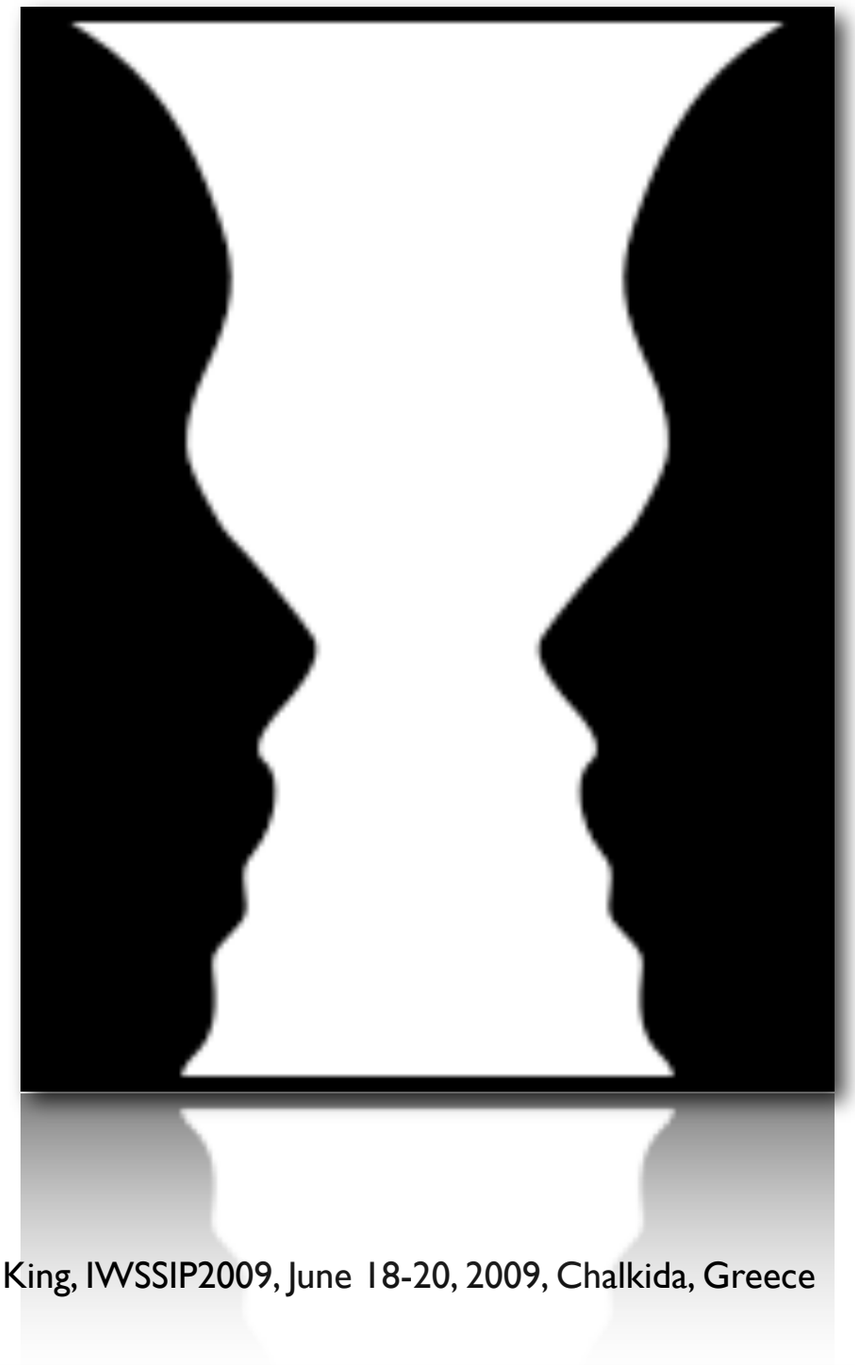
Web 2.0

- Web as a medium vs. **Web as a platform**
- Read-Only Web vs. **Read-and-Write Web**
- Static vs. **Dynamic**
- Restrictive vs. **Freedom & Empowerment**
- Technology-centric vs. **User-centric**
- Limited vs. **Rich User Experience**
- Individualistic vs. **Group/Collective Behavior**
- Consumer vs. **Producer**
- Transactional vs. **Relational**
- Top-down vs. **Bottom-up**
- People-to-Machine vs. **People-to-People**
- Search & browse vs. **Publish & Subscribe**
- Closed application vs. **Service-oriented Services**
- Functionality vs. **Utility**
- Data vs. **Value**

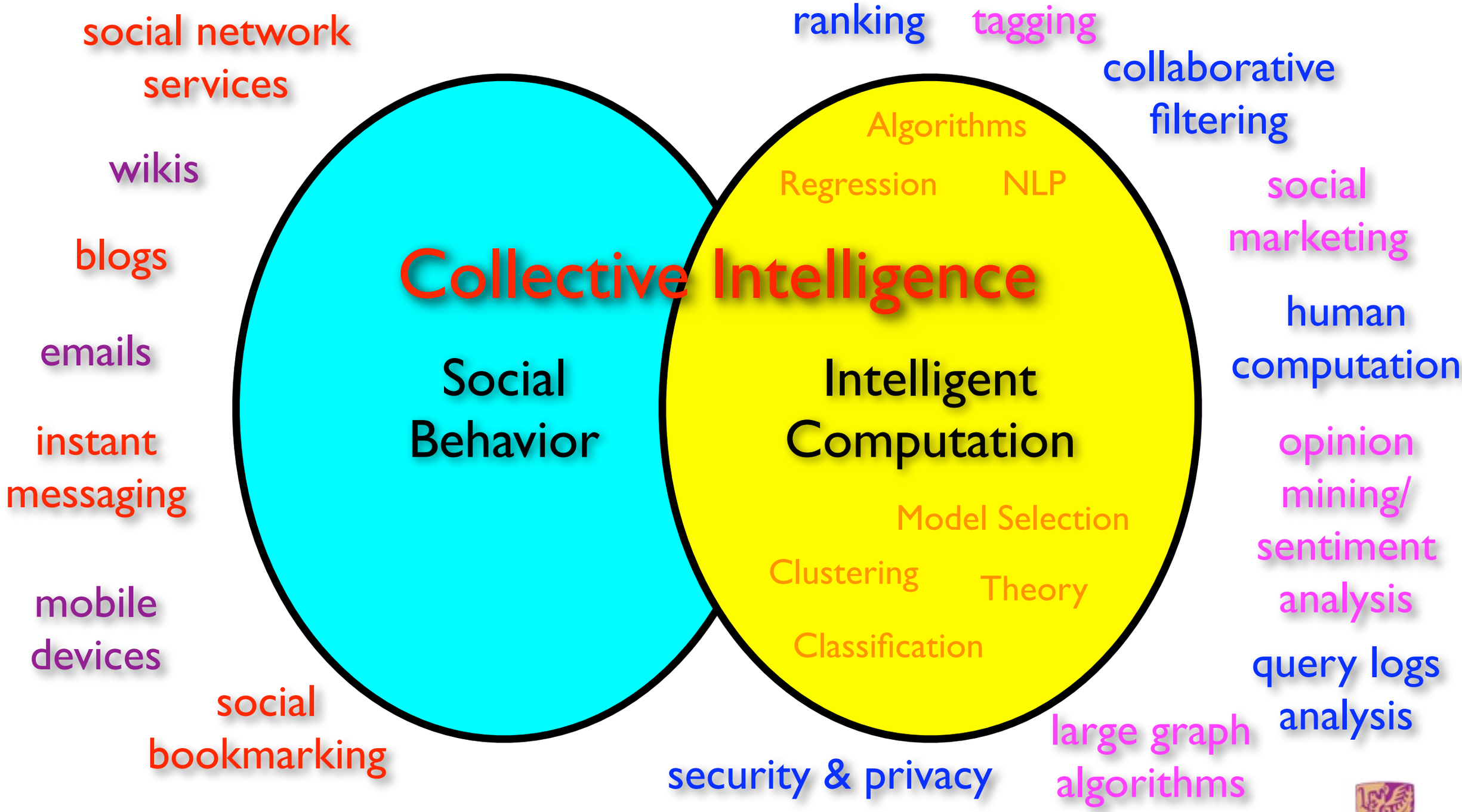
The New Paradigm Shift: The Emergence of Social Computing, Irwin King, IWSSIP2009, June 18-20, 2009, Chalkida, Greece

Web 2.0 Revolution

- Glocalization-think globally and act locally!
- Weblication-Web is the application!
- Three C's
  - Connectivity
  - Collaboration
  - Communities



Social Computing



The New Paradigm Shift: The Emergence of Social Computing, Irwin King, IWSSIP2009, June 18-20, 2009, Chalkida, Greece

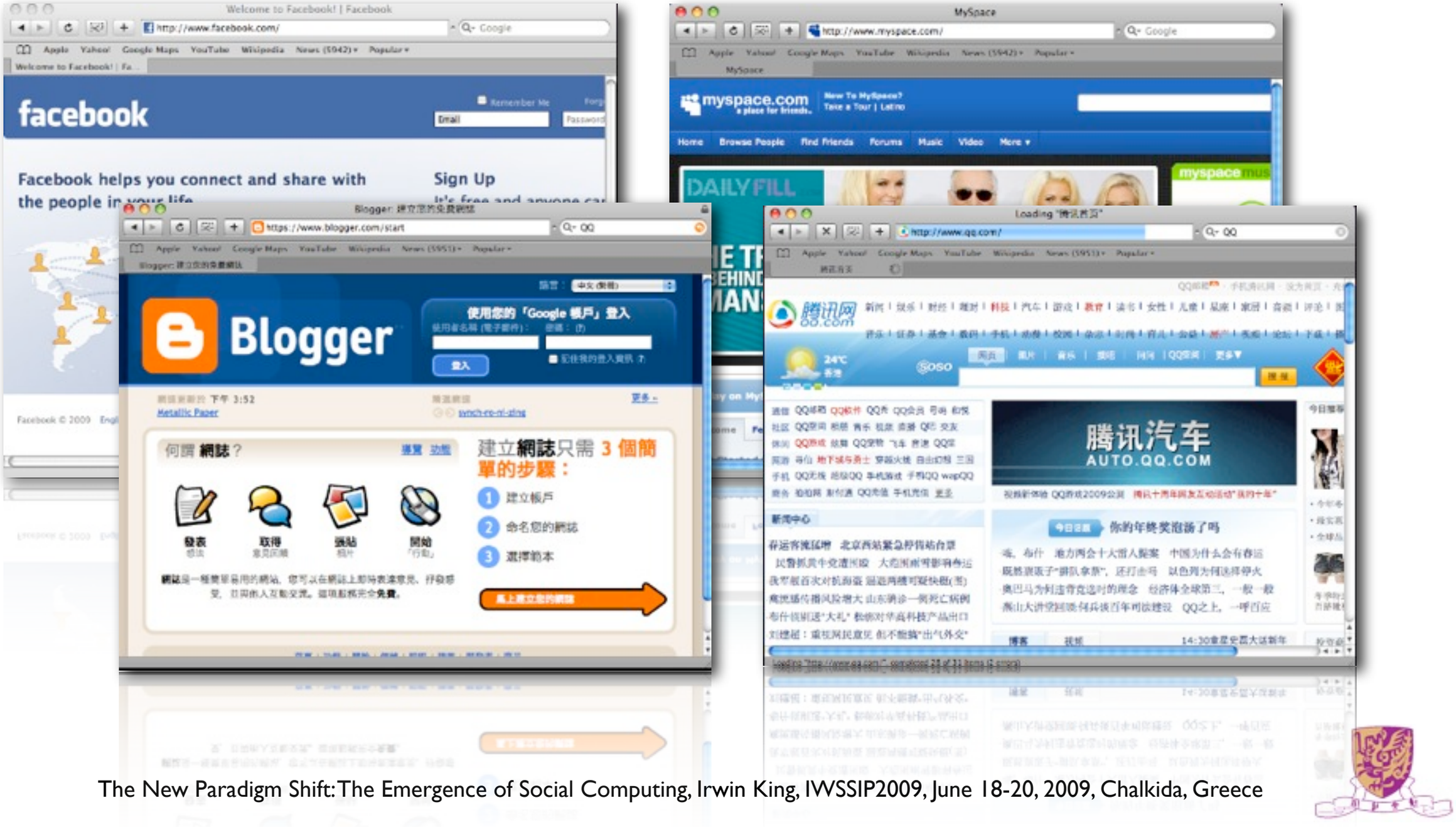
Definition of Social Computing

- Any Computer-mediated communication and interaction
- In the weaker sense: supporting any sort of social behavior
  - blogs, email, instant messaging, wiki, social network services, social bookmarking
- In the stronger sense: supporting "computations" that are carried out by a group of people
  - recommender systems, online auctions, prediction markets, reputation systems, tagging, verification games

Social Networking in Greece

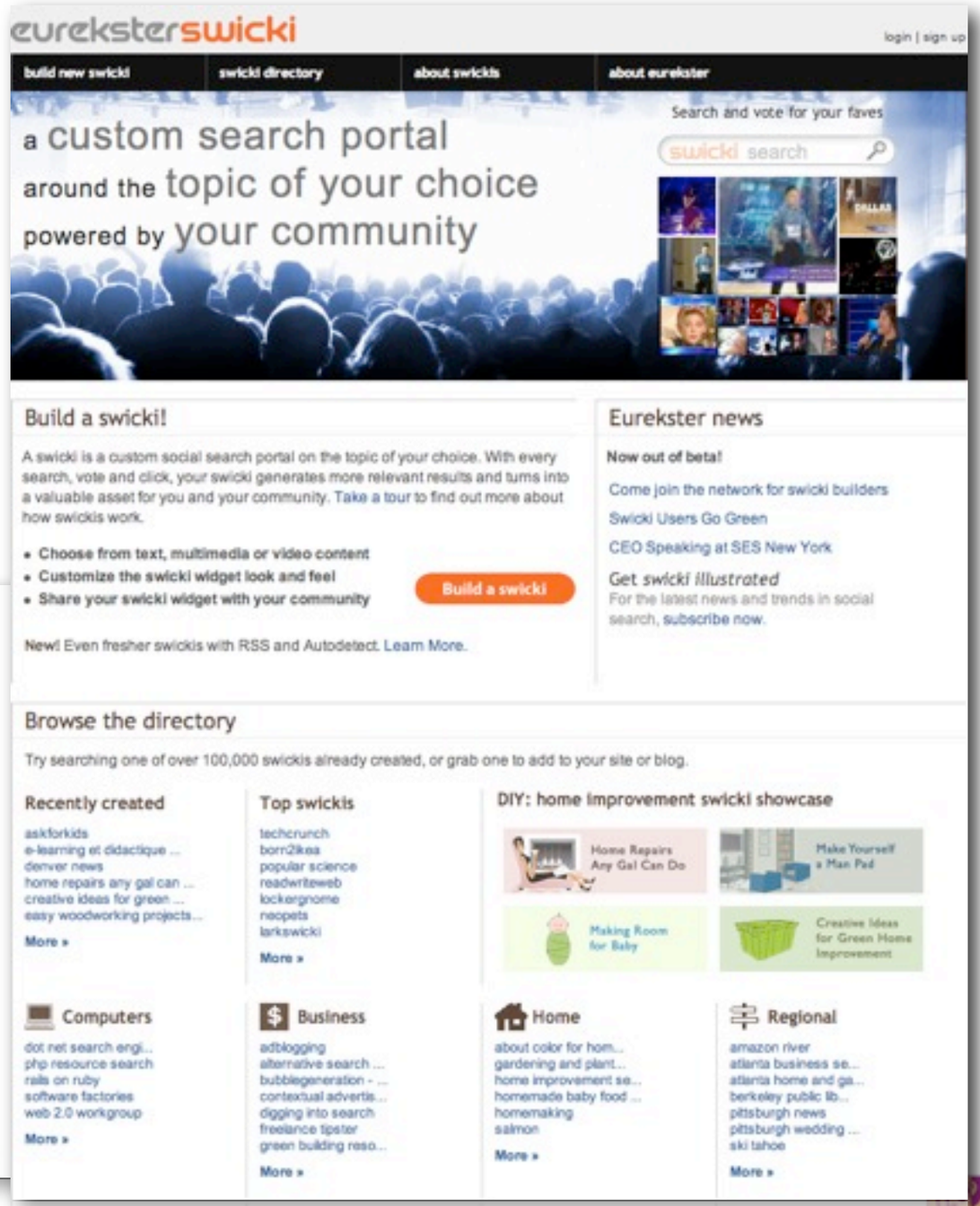
Social Networking Sites

- Example of Social Networking Sites: FaceBook, MySpace, Blogger, QQ, etc.

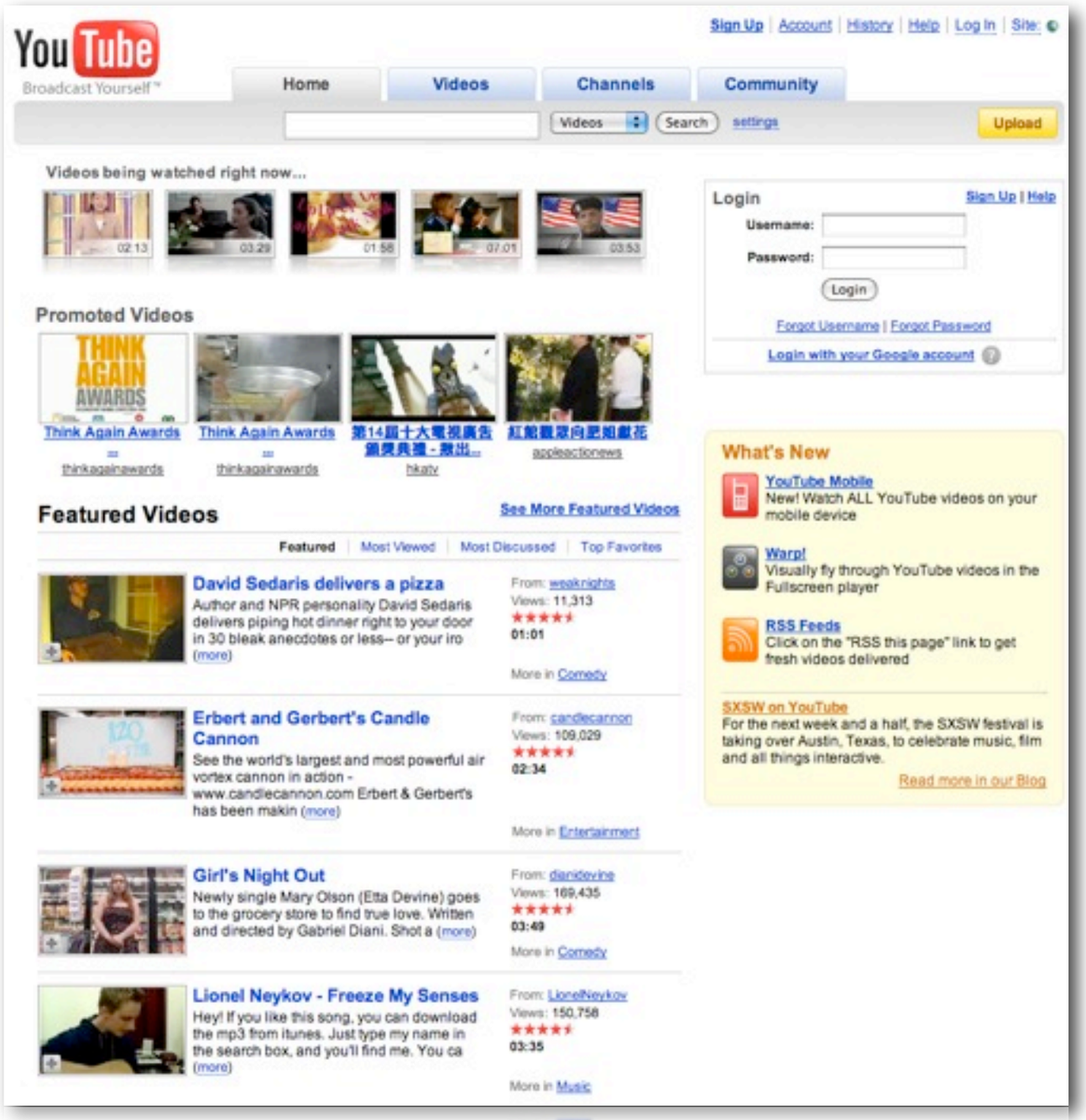


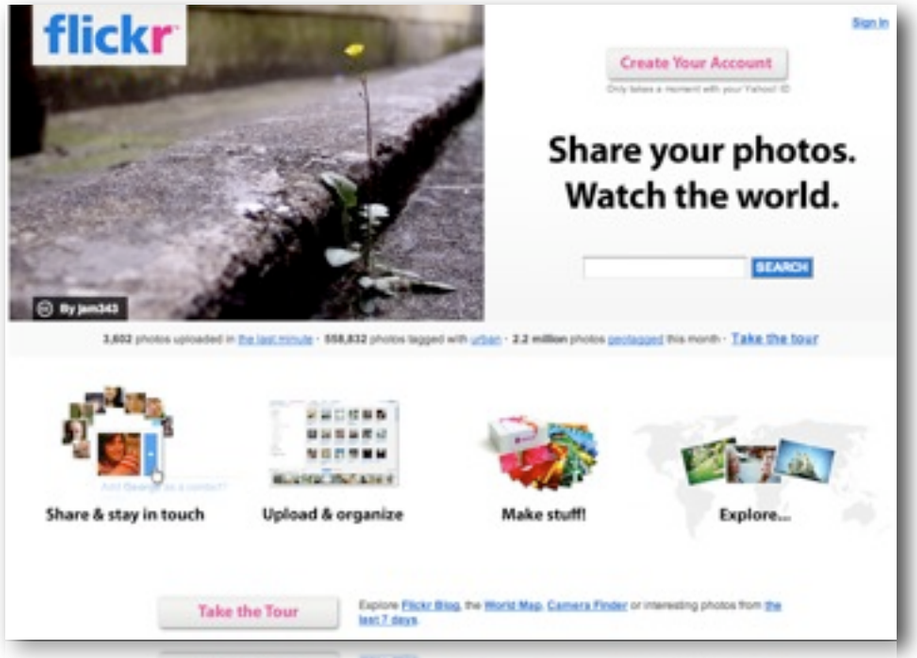
Social Search

- Social Search Engine
- Leveraging your social networks for searching

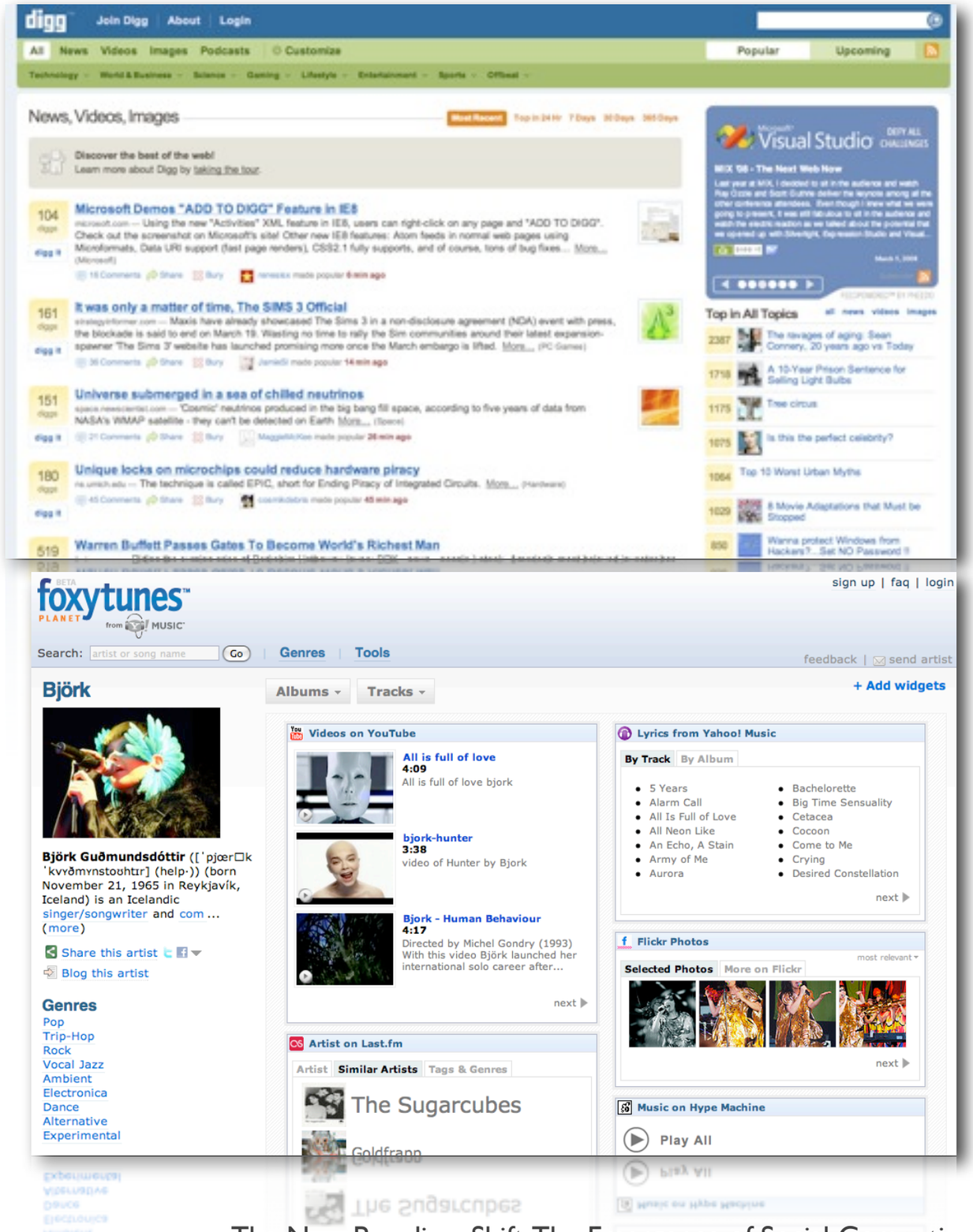


Social Media

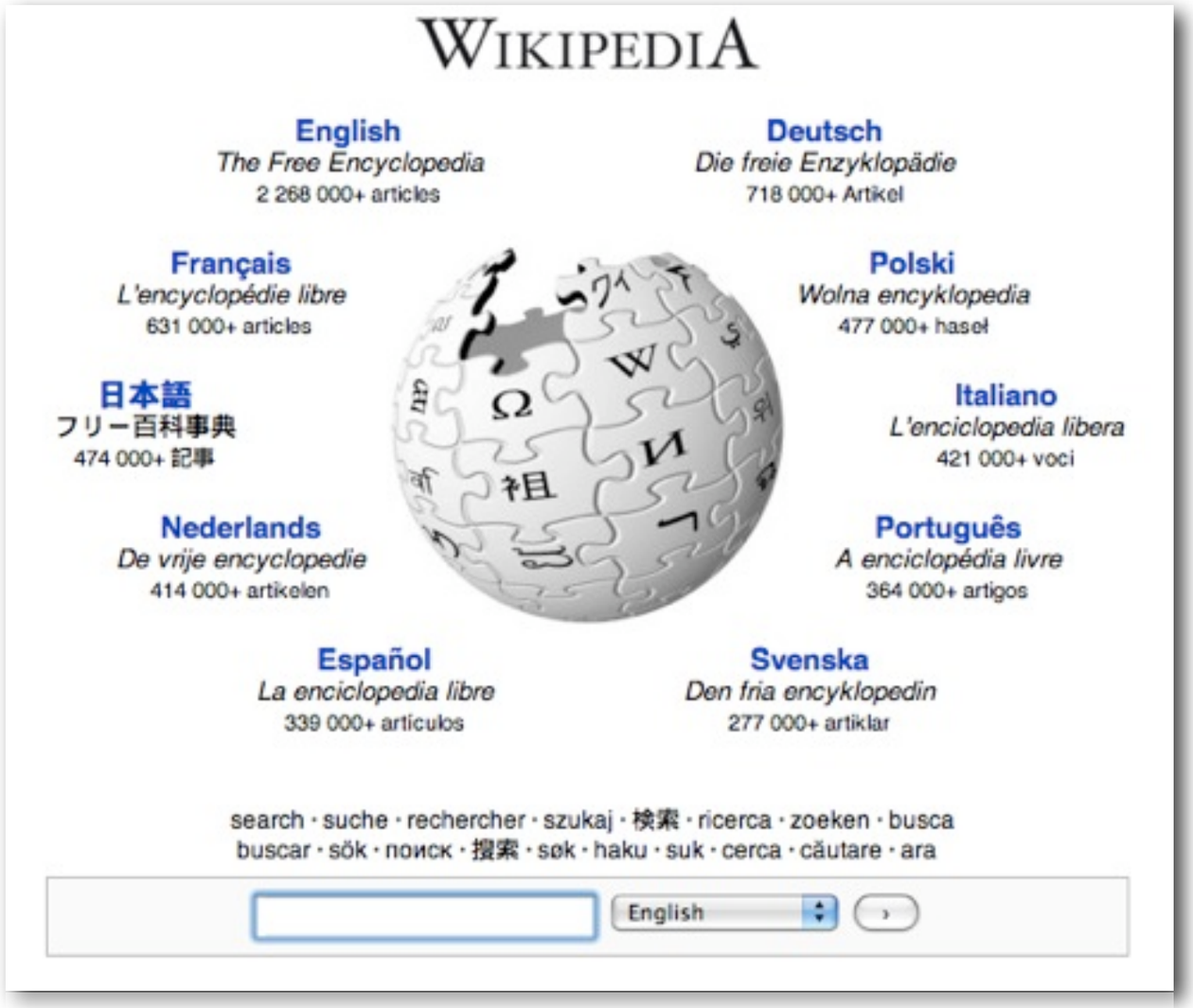




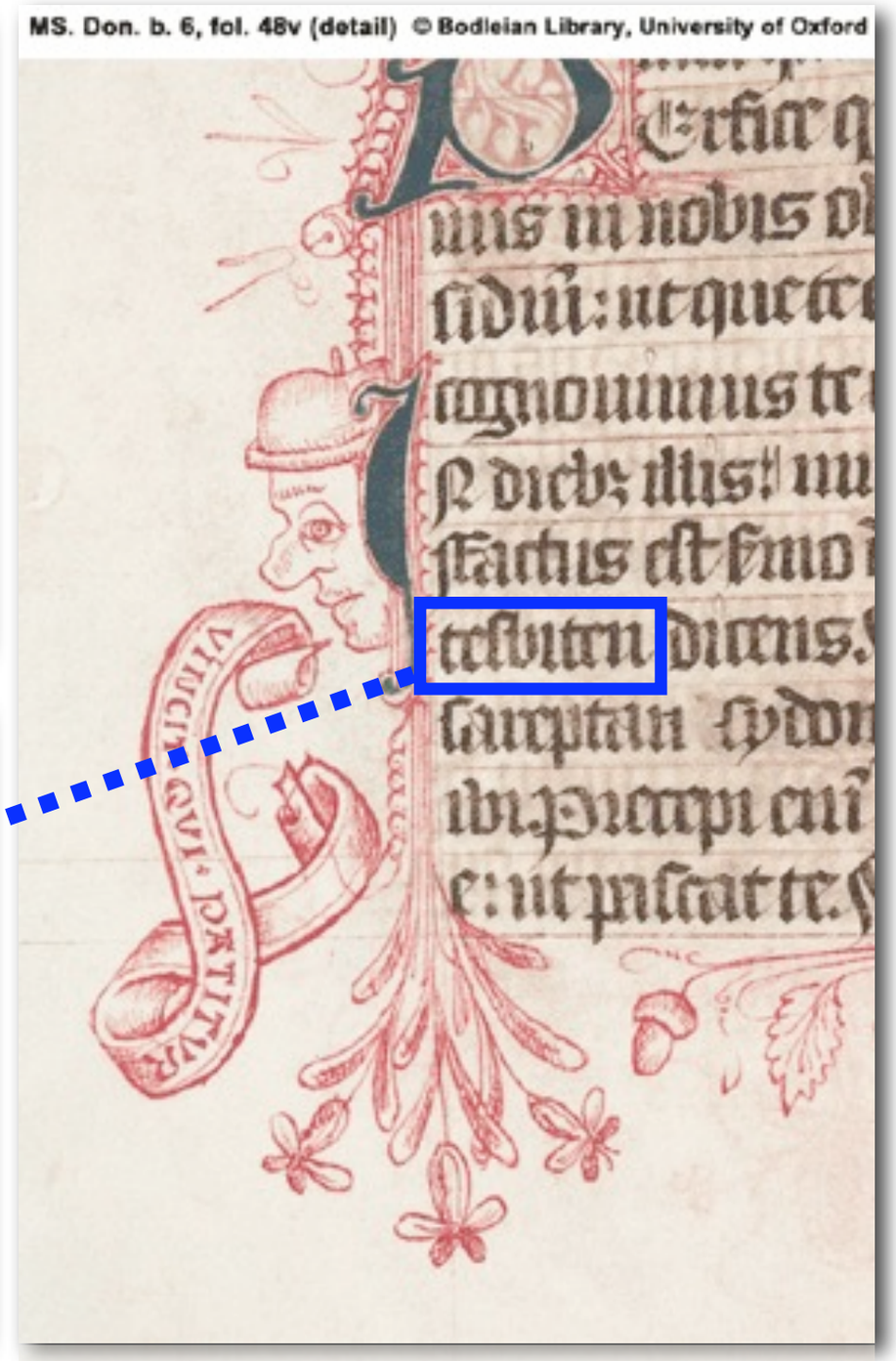
Social News/Mash Up



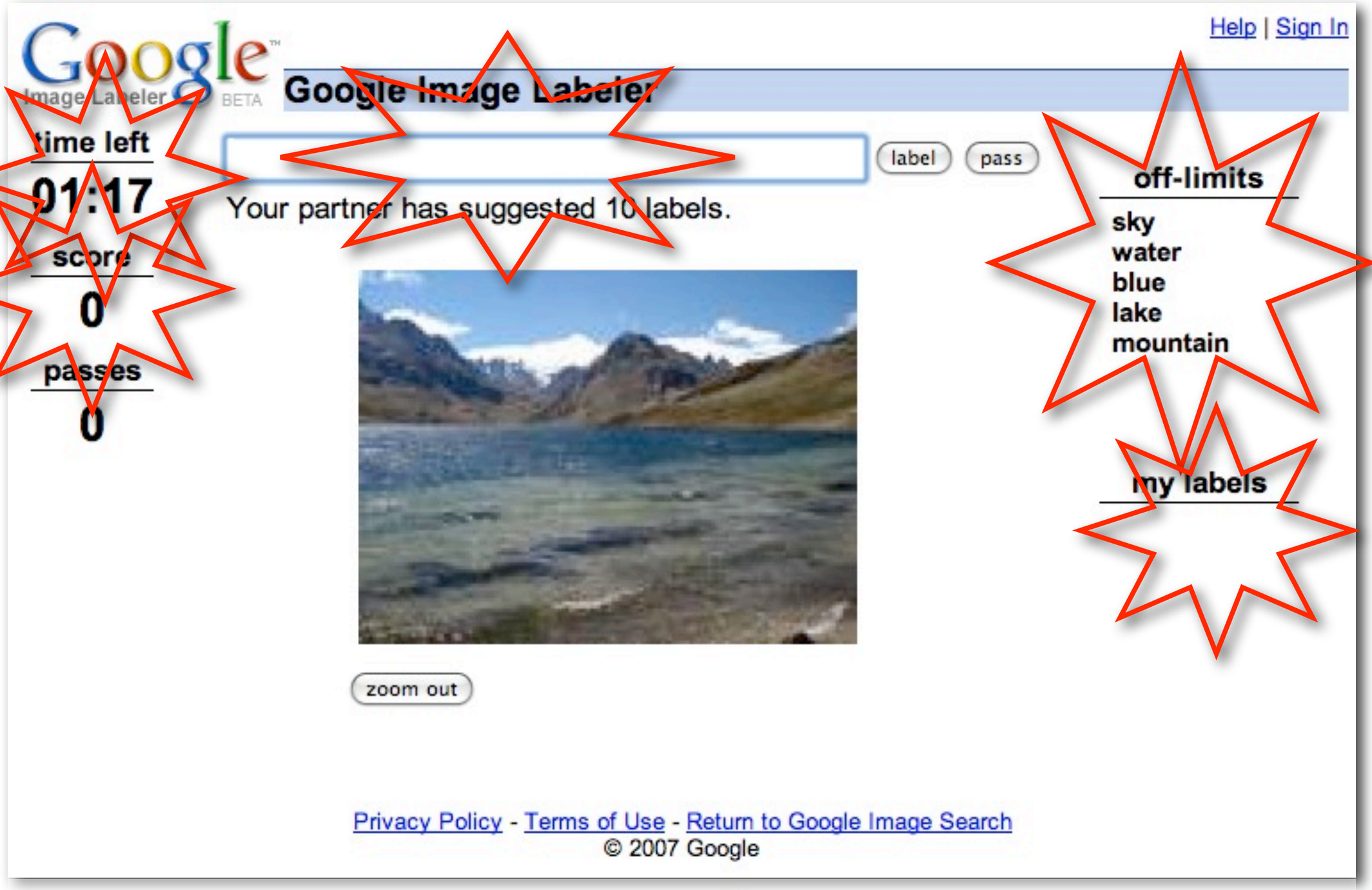
Social Knowledge Sharing



Social/Human Computation

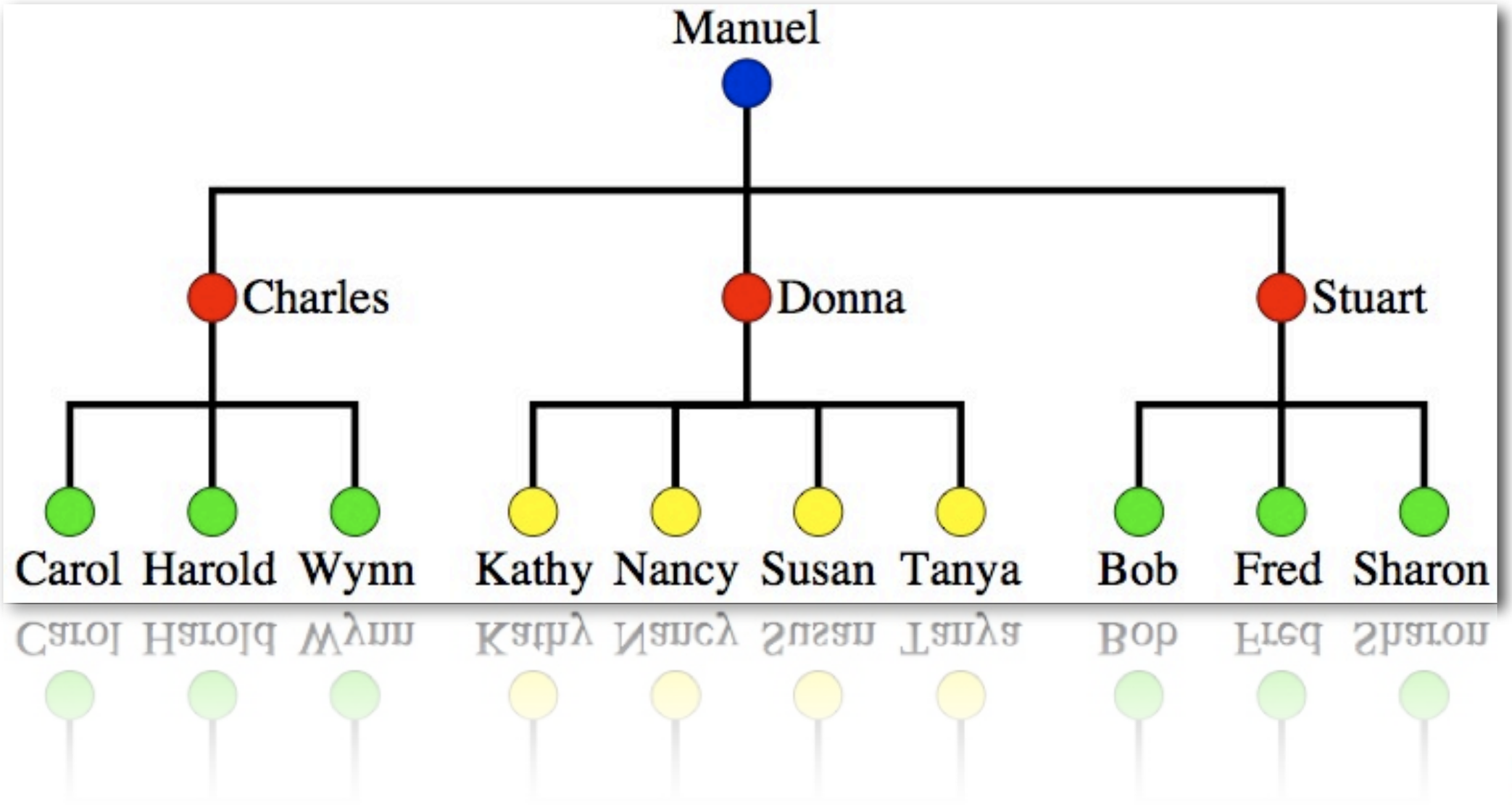


Human Computation

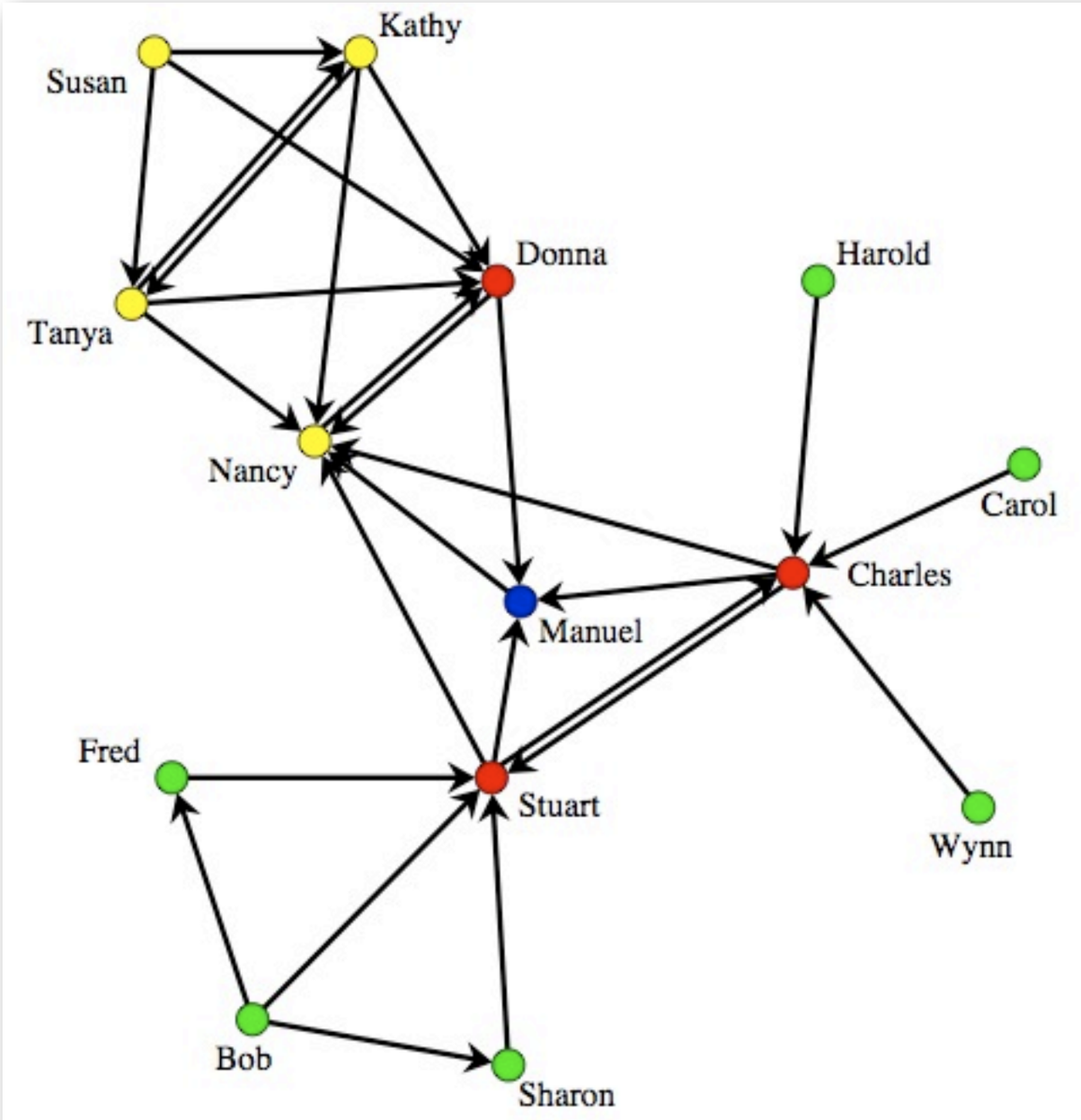


The Social Web

Organizational Chart



Social Network Chart



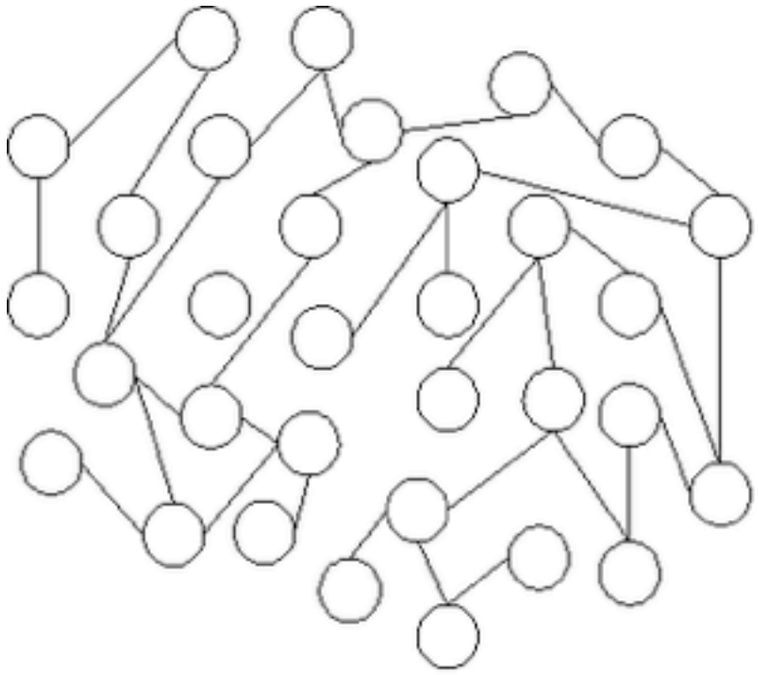
1.00	Nancy	(secretary)
0.66	Donna	(supervisor)
0.57	Manuel	(manager)
0.19	Stuart	(supervisor)
0.17	Charles	(supervisor)
0.08	Kathy	(secretary)
	Tanya	(secretary)
0.02	Fred	(auditor)
	Sharon	(auditor)
0.00	Bob	(auditor)
	Carol	(auditor)
	Harold	(auditor)
	Wynn	(auditor)
	Susan	(secretary)

Characteristics of Social Networks

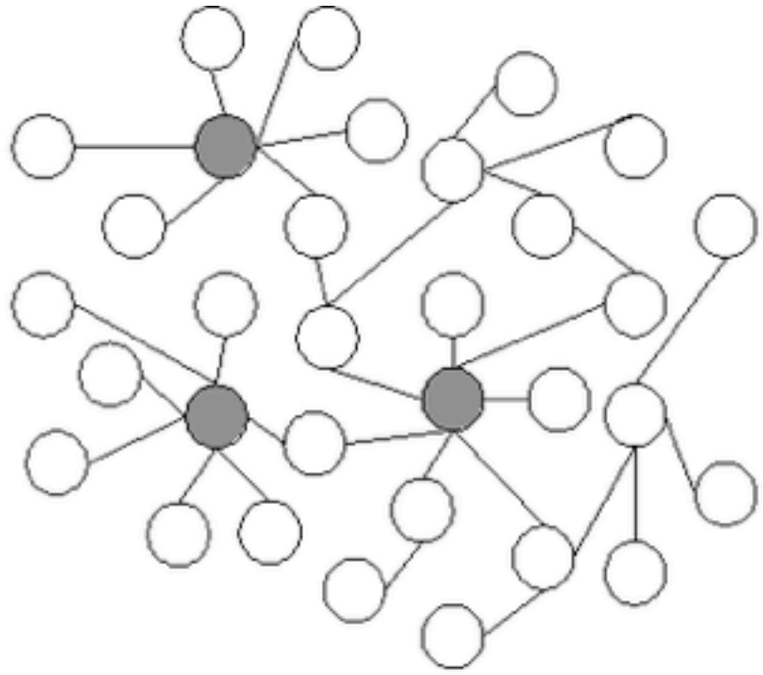
- Power-law distribution and small world network characteristics of
  - Co-authorship network
  - Email network
  - World Wide Web
  - Blog network
  - Instant-messaging network
  - Mobile network

Scale-free Networks

- A scale-free networks is a network whose degree distribution follows the Power Law
  - With a few highly connected nodes that serve as hubs and many nodes with only a few connections



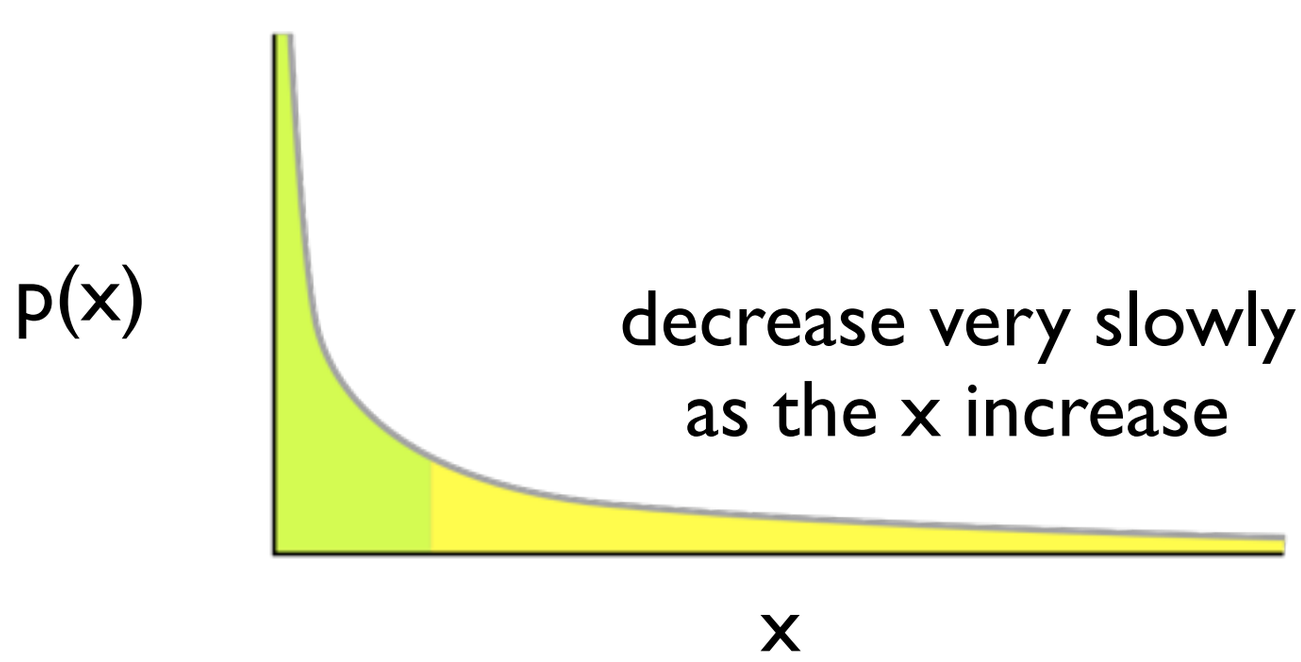
(a) Random network



(b) Scale-free network

Power-law distribution

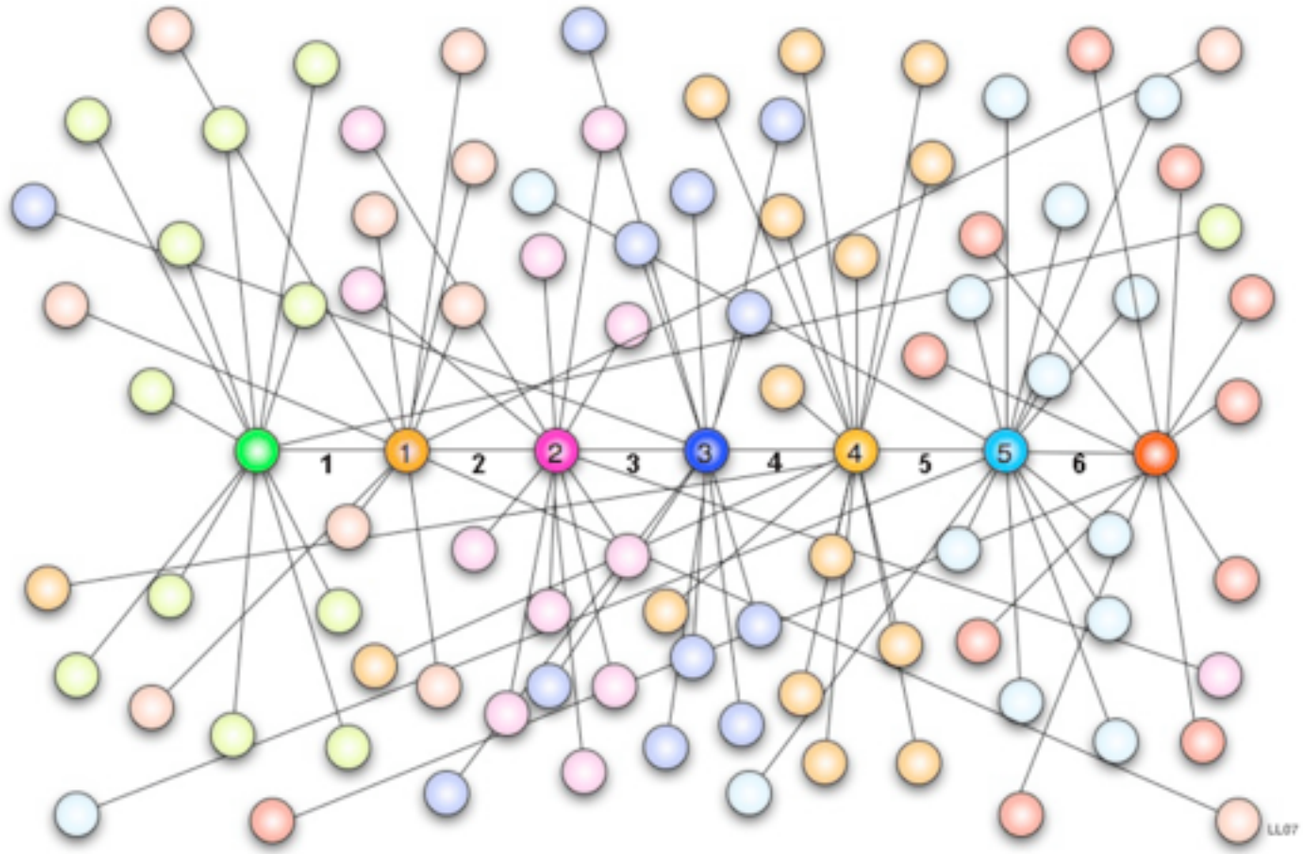
- The Power Law: mathematical relationship between two quantities: $p(x) \propto x^{-\lambda}$
  - p(x) is the probability to encounter value x
  - λ is the exponent of the power law



An example power law graph with a long tail, to the left are the few that dominate (also known as the 80-20 rule)

Small World Network

- A small world network is one in which most nodes are not connected to each other and yet the average path between most nodes is relatively short



Six degree of separation

Experiments on Co-authorship Networks

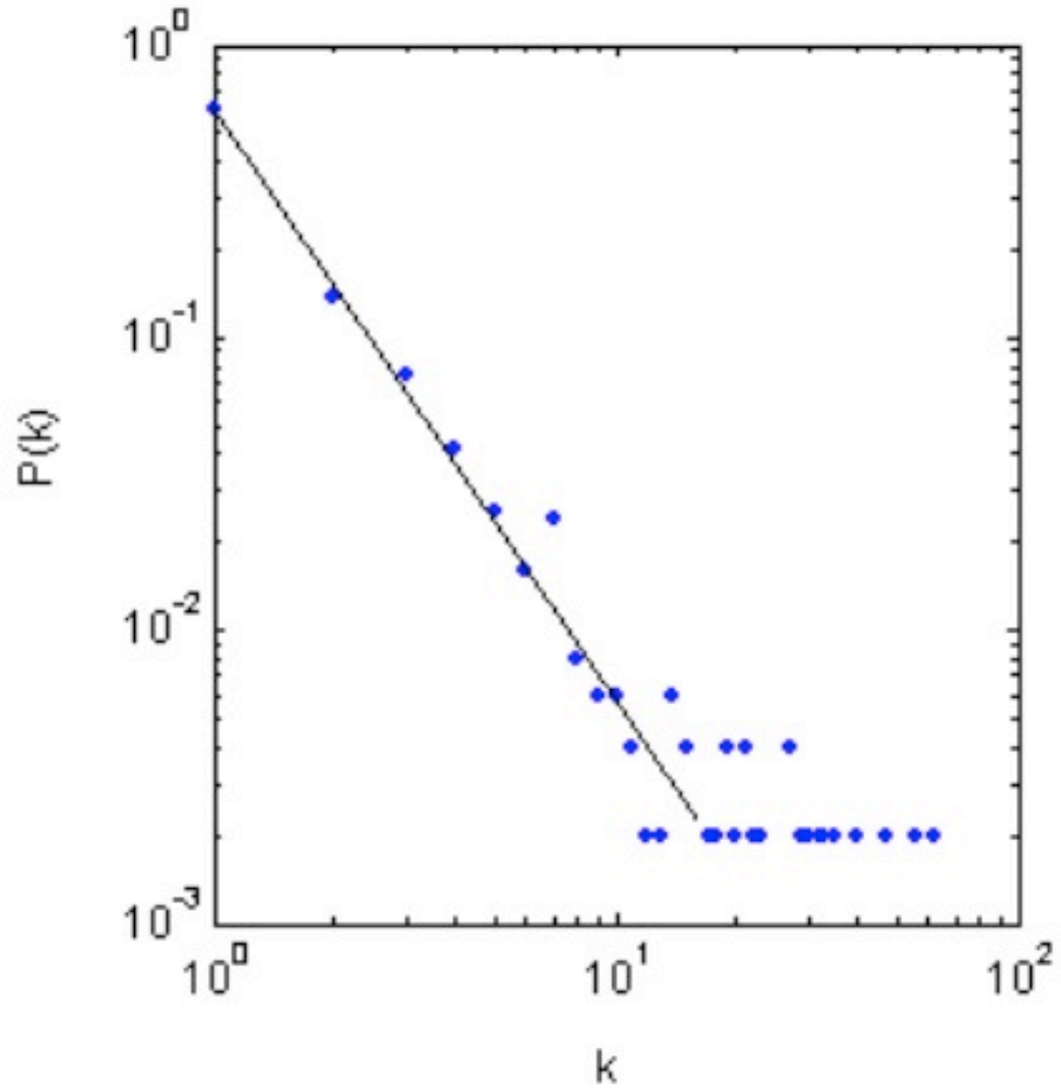
[Chunguang Li]

- Erdös's collaboration network, which is a star-shaped network with Erdös at the centre.
  - Paul Erdös, a great mathematician who passed on September 20, 1997, at the age of 83
  - He published more than 1,600 mathematical research papers in his lifetime.
  - He had 507 coauthors, in which there were 306 coauthors who published 1 paper with him, and in which the coauthor having the largest number of joint publications with him was Sarkozy, 62 papers

Experiments on Co-authorship Networks

[Chunguang Li]

$$P(k) = k^{-\lambda}, \lambda = 2.0 \pm 0.2$$



The distribution of connection strengths between Erdös and his coauthors

Experiments on Co-authorship Networks

[A.L. Barabasi, 2002]

- The databases contain article titles and authors of all relevant journals in the field of Mathematics (M) and Neuro-Science (NS), published in the period 1991–98

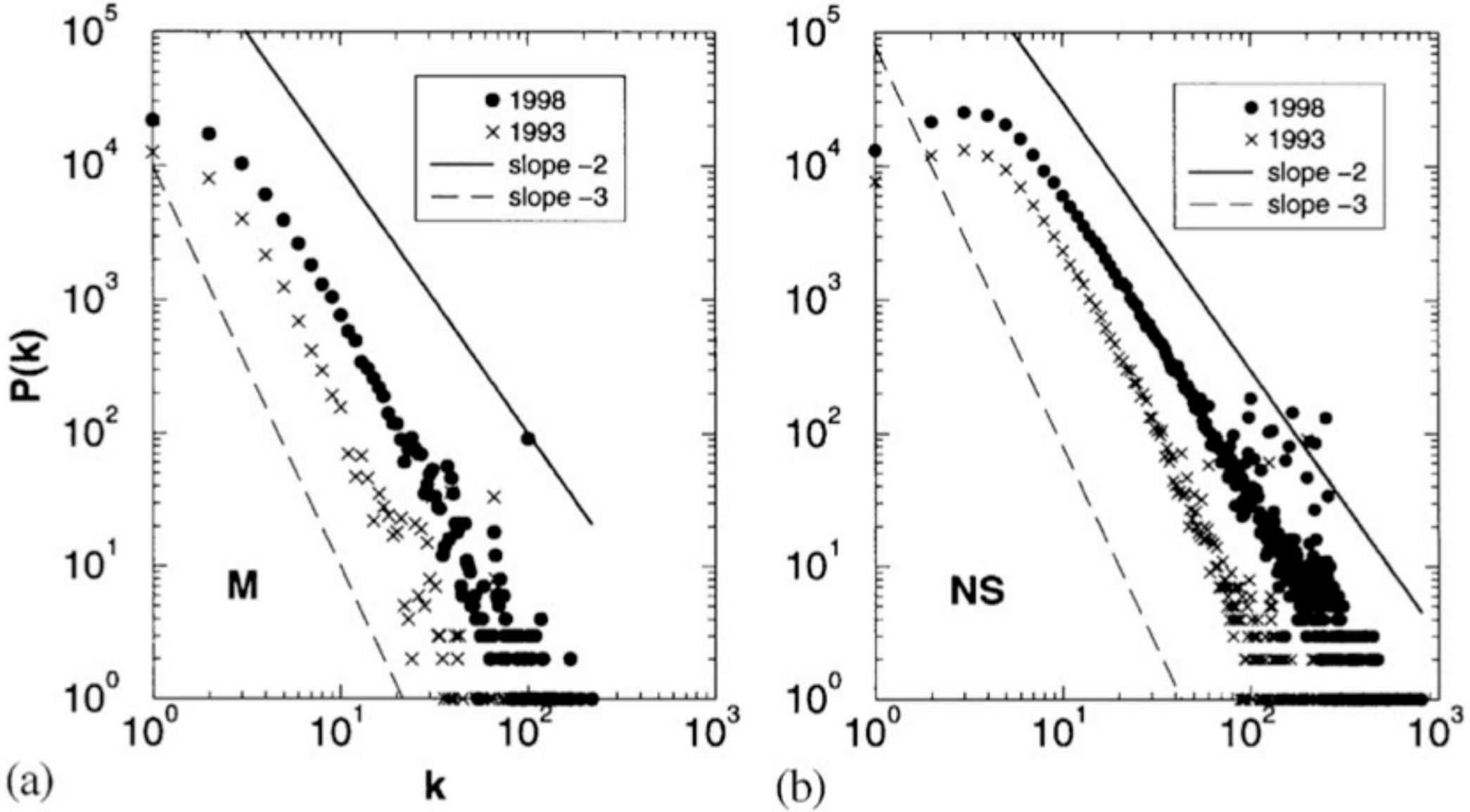
- In mathematics our database contains 70,975 different authors and 70,901 papers

- In NS the number of different authors is 209,293 and the number of published papers is 210,750

Experiments on Co-authorship Networks

[A.L. Barabasi, 2002]

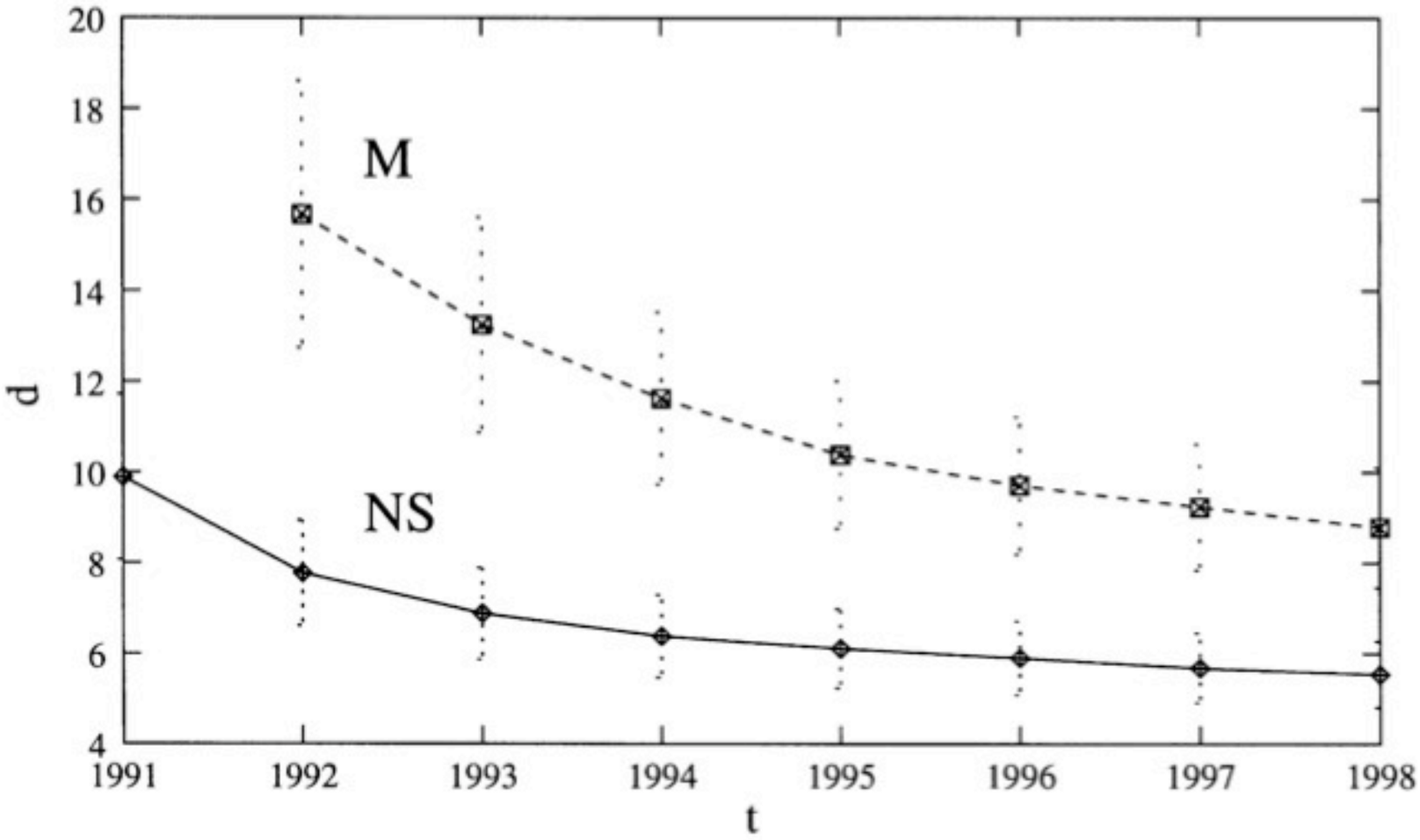
$\lambda_M = 2.4$ $\lambda_{NS} = 2.1$



Degree distribution for both M and NS databases

Experiments on Co-authorship Networks

[A.L. Barabasi, 2002]



Average separation in the M and NS databases

Experiments on Email Network

[Duncan Watts, 2001]

- Recreate Milgram's experiment on the internet
- An e-mail message as the "package" that needed to be delivered
- 48,000 senders and 19 targets (in 157 countries)
- Watts found that the average (though not maximum) number of intermediaries was around 6

Experiments on Email Network

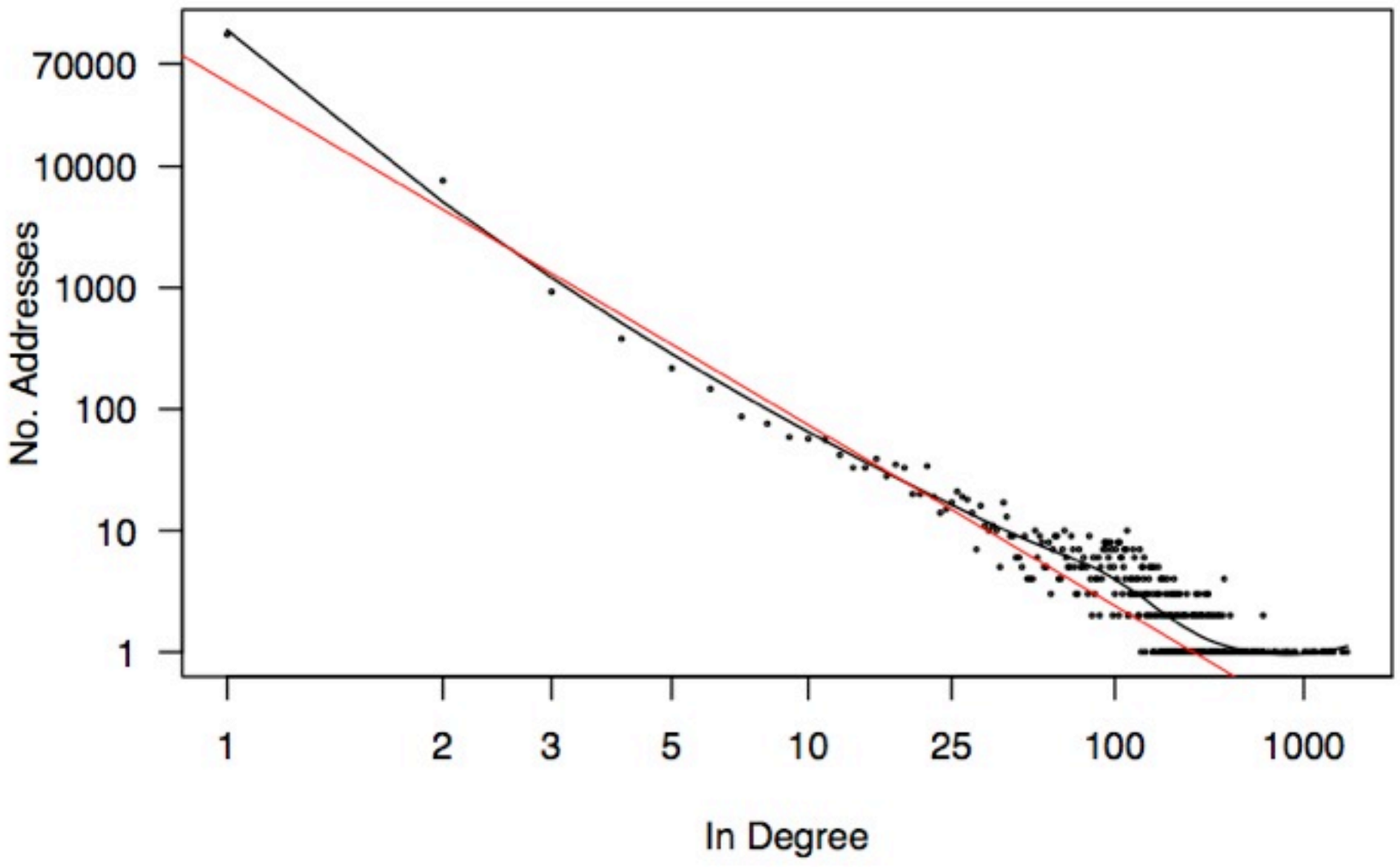
[Diane Lambert, 2004]

- A complete set of email records obtained from several servers over a 75 day period with delivery dates on or after September 29, 2004
  - 805,100 messages correspond to 176,761 addresses
  - A total of 303,499 unique directed links and 296,742 undirected links
  - A directed link is created between addresses A and B if A sends one or more messages to B during the study and an undirected link if A sends B email or vice versa.

Experiments on Email Network

[Diane Lambert, 2004]

$$p(k) \propto (\log(k+1))^{-5.9}$$

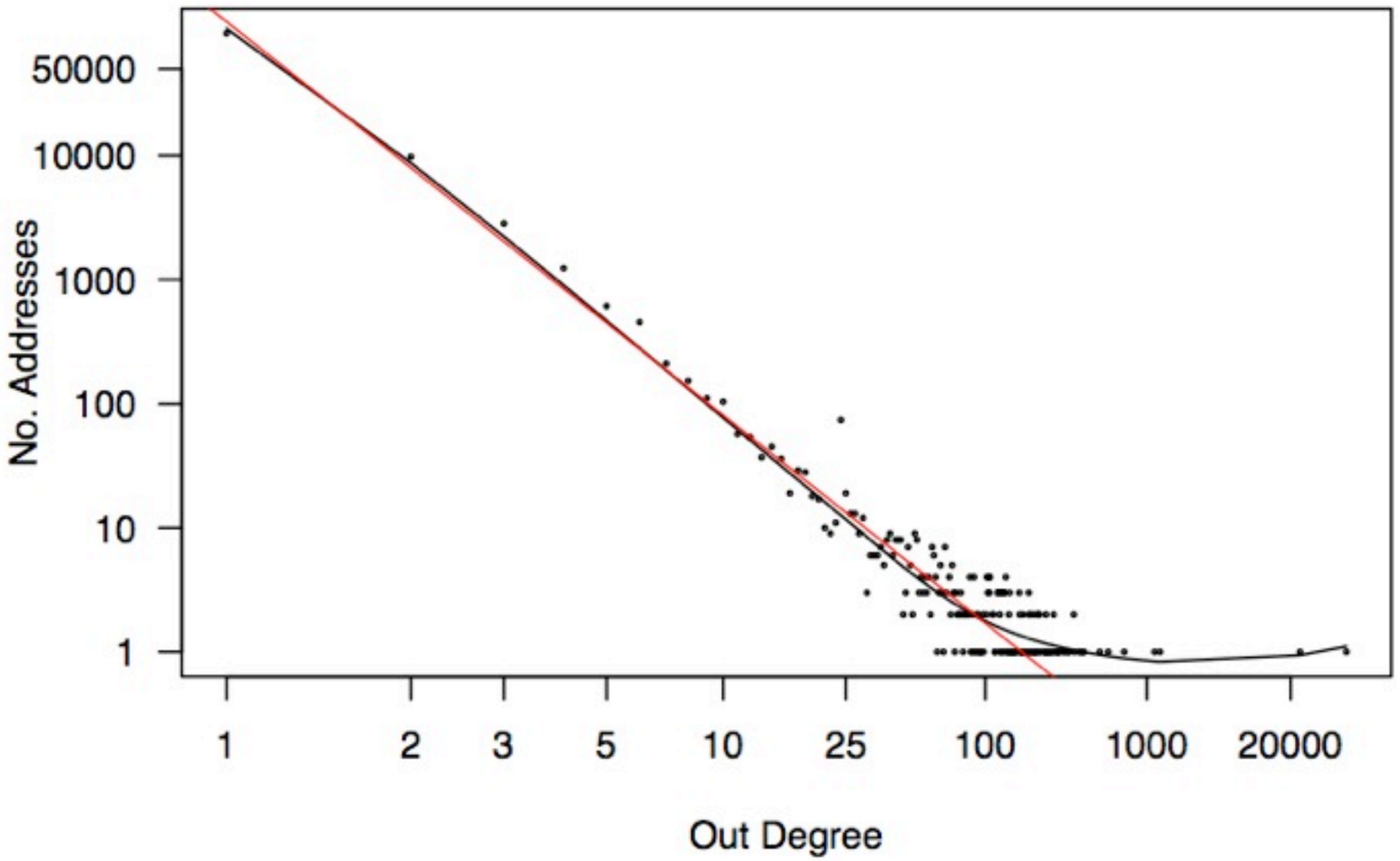


The in-degree distribution

Experiments on Email Network

[Diane Lambert, 2004]

$$p(k) \propto (\log(k+1))^{-5.2}$$

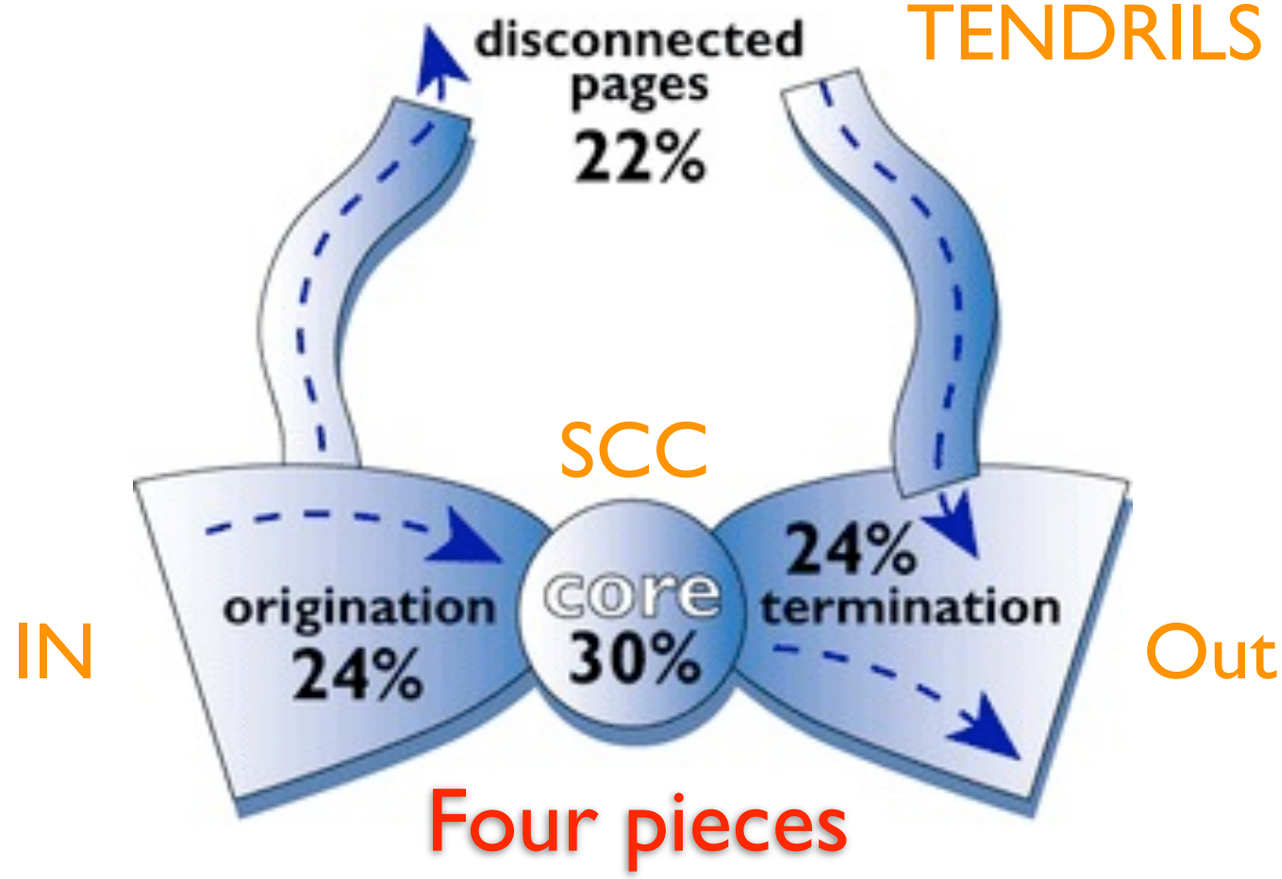


The out-degree distribution

Experiments on Web Crawls

[Andrei Broder, 2000]

- A web crawl (breadth-first search) of approximately 200 million pages and 1.5 billion links (2000)
- If links are treated as undirected edges, over 90% nodes form a single connected component



Four pieces

Experiments on Web CrawIs

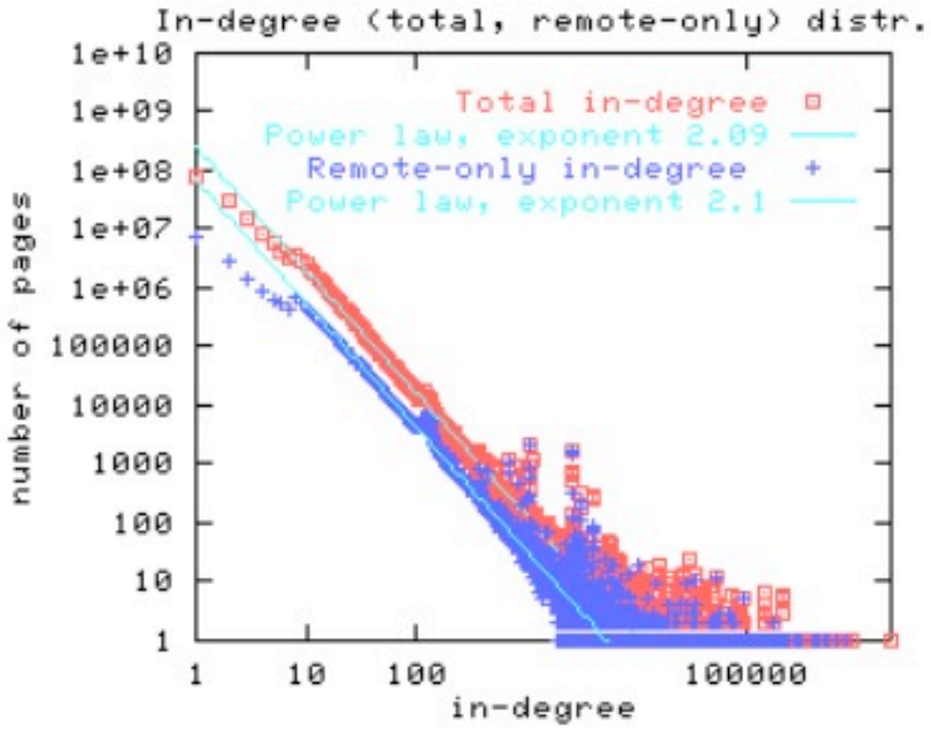
[Andrei Broder, 2000]

- Results
  - Path exists from random source to destination is 24%
  - If a directed path exists, its average will be about 16
  - If an undirected path exists, its average will be about 6

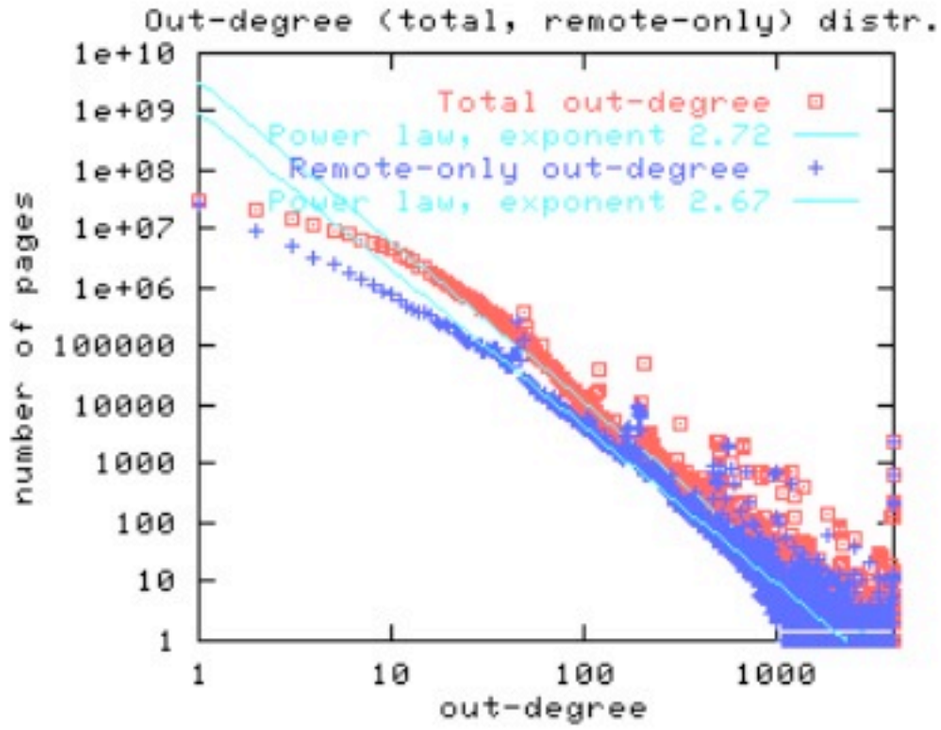
Experiments on Web Crawls

[Andrei Broder, 2000]

$$N(i) \propto i^{-2.1}$$



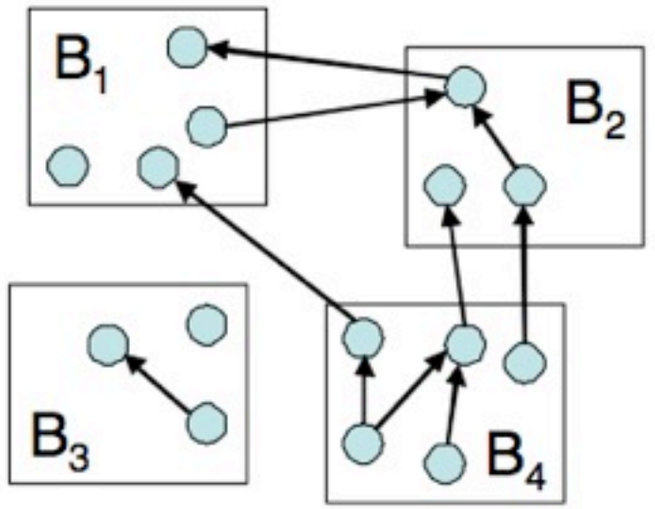
$$p(o) \propto o^{-2.7}$$



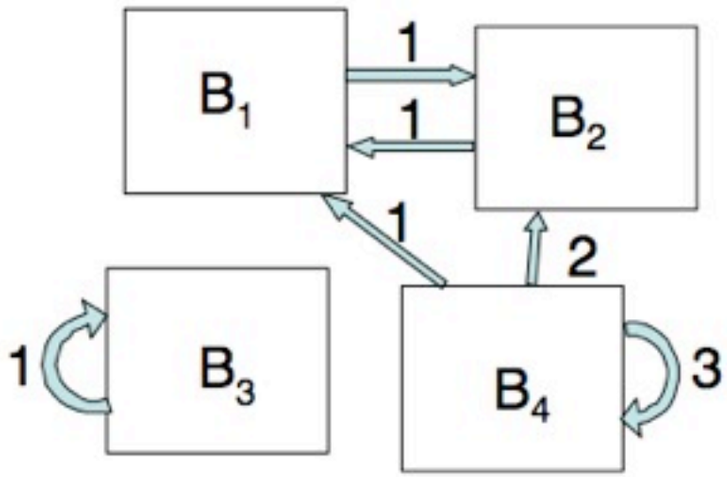
In-degree and out-degree distributions subscribe to power law.
The Law also holds if remote-only edges are considered

Experiments on Blog Graph

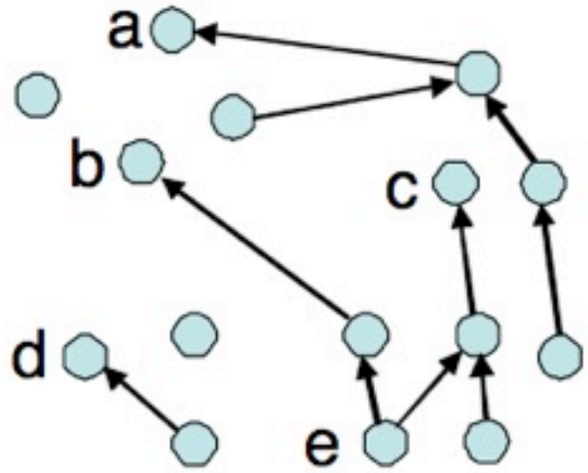
[Jury Leskovec, 2007]



(a) Blogosphere



(b) Blog network



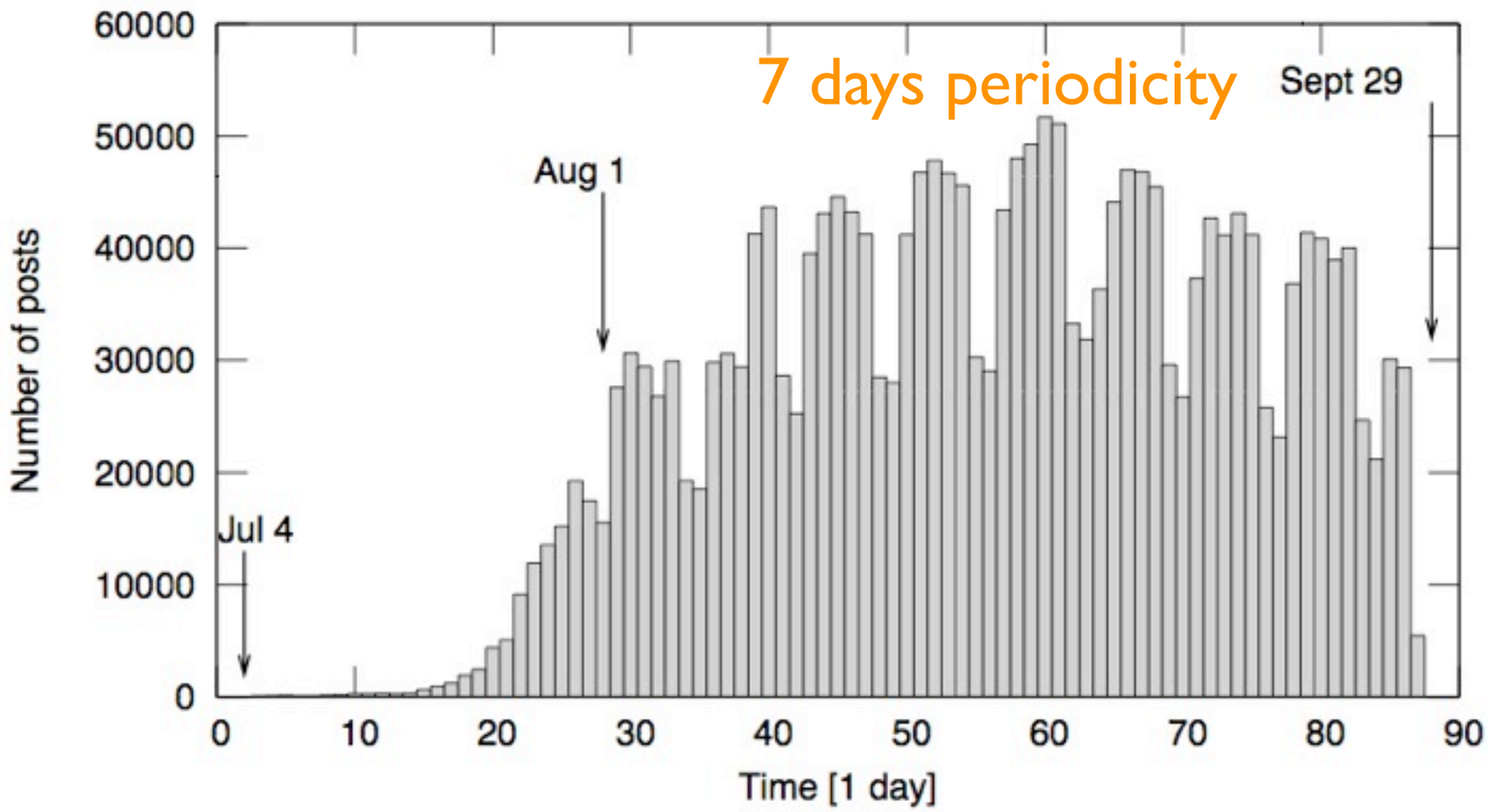
(c) Post network

The model of the blogosphere. Squares represent blogs and circles blog-posts

Experiments on Blog Graph

[Jury Leskovec, 2007]

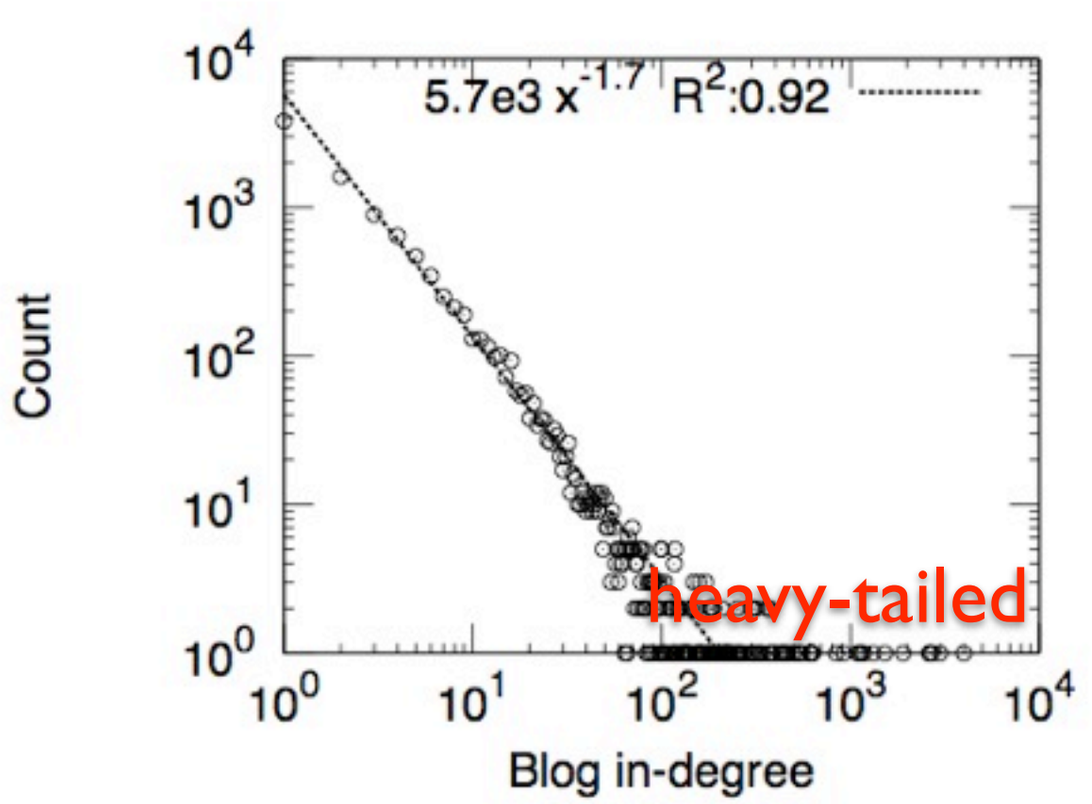
- A larger set of 21.3 million posts from 2.5 million blogs
- Extract 2.4 million posts in 44 thousand blogs from Aug. and Sept. 2005 (90 days)

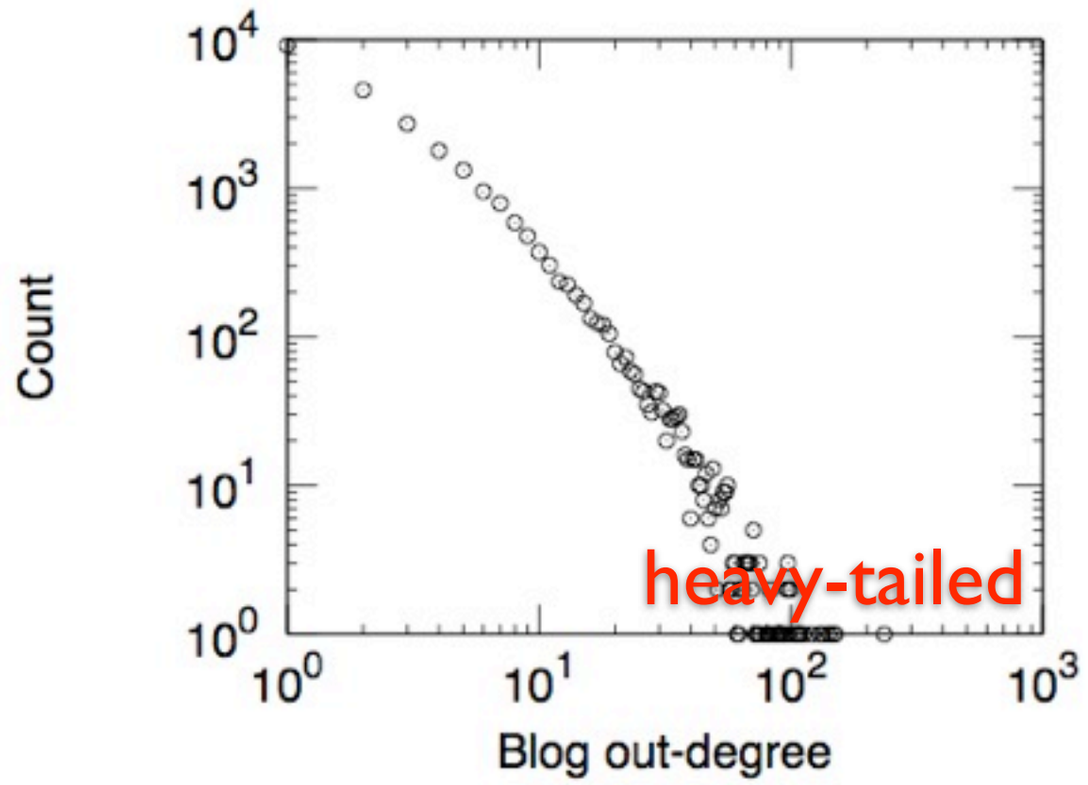


Experiments on Blog Graph

[Jury Leskovec, 2007]

In-degree distribution has a very shallow power-law (-1.7), which suggests strongly rich-get-richer phenomena

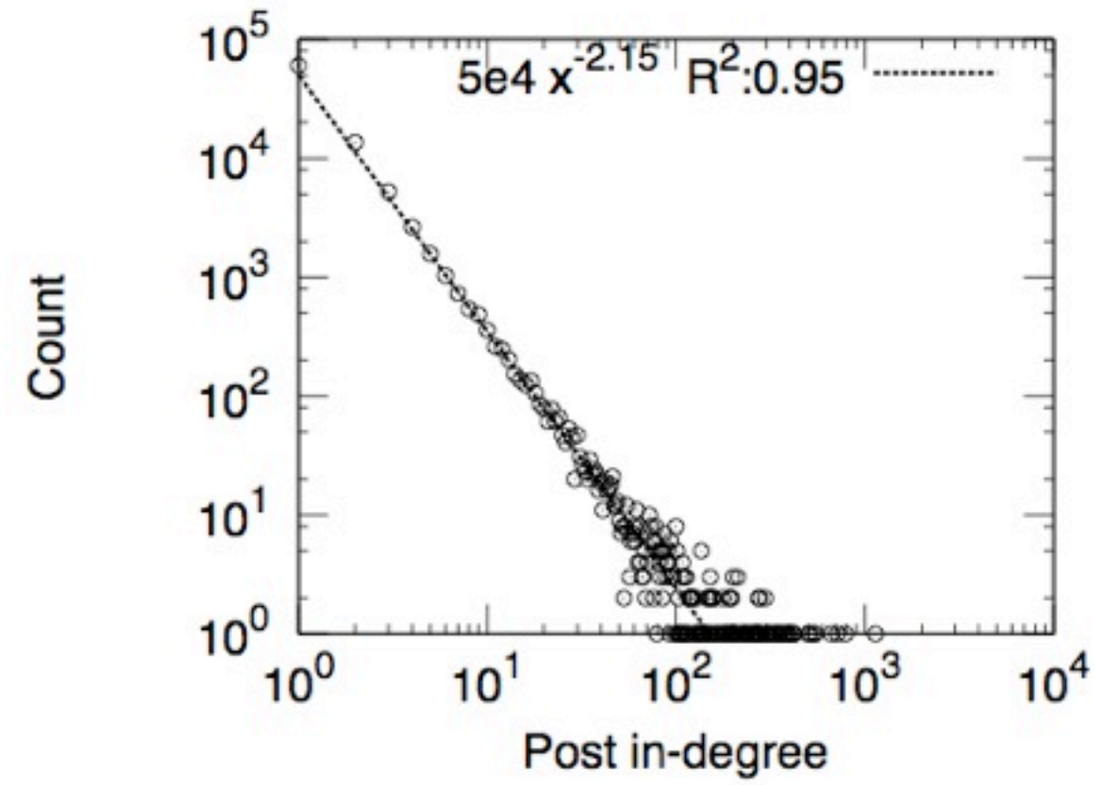


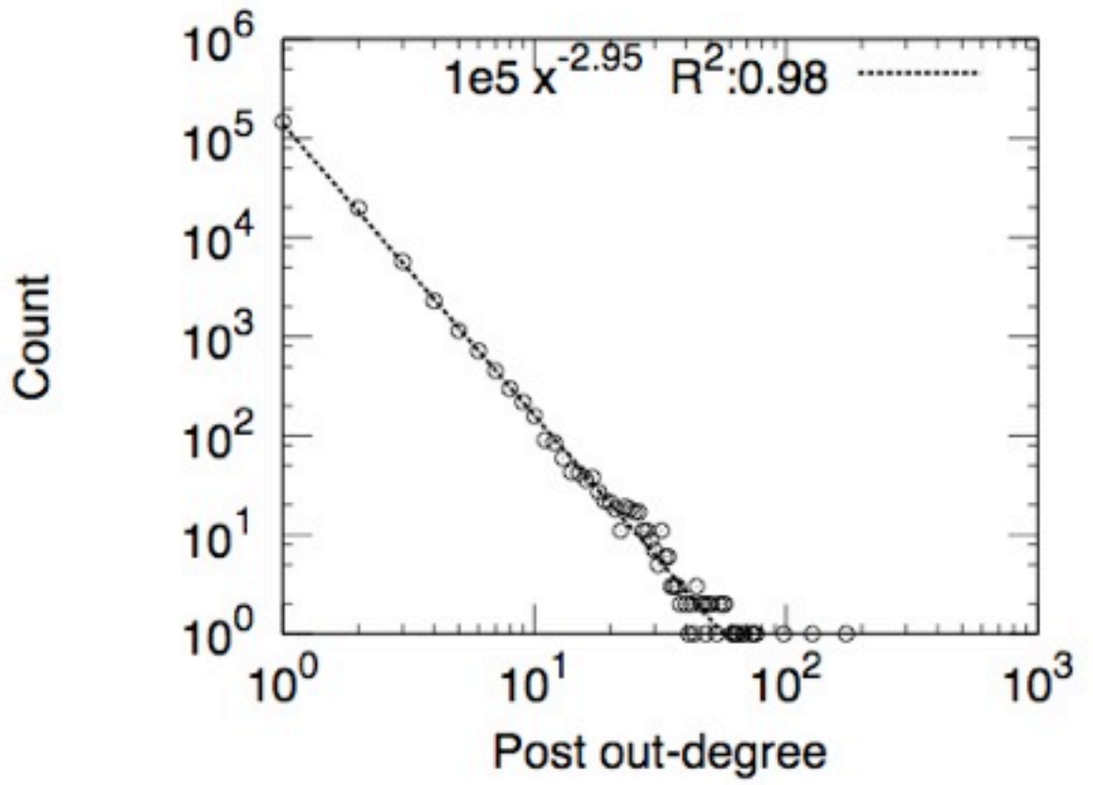


In- and out-degree distribution of the Blog network

Experiments on Blog Graph

[Jury Leskovec, 2007]





Post network in- and out-degree distribution

Experiments on MSN

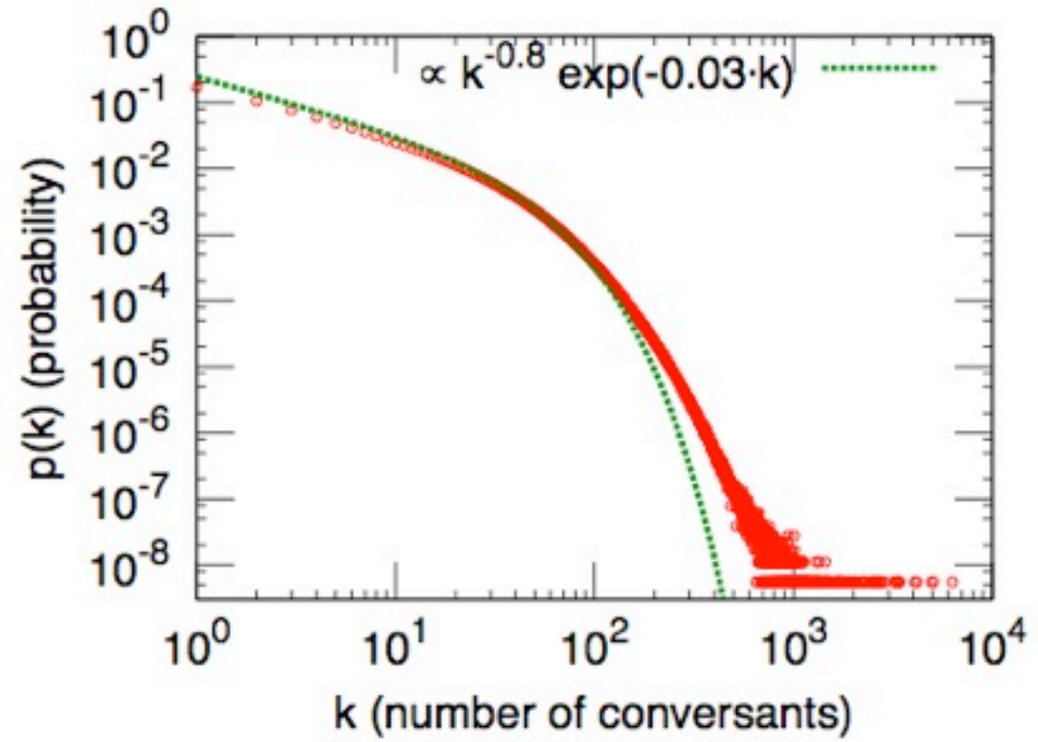
[Jury Leskovec, 2008]

- A study of anonymized data capturing a month of communication activities within the whole of the Microsoft Messenger instant-messaging system
  - Dataset: 240 million people with 30 billion conversation
  - Graph: 180 million nodes and 1.3 billion undirected edges (only connected people who are buddies and have communicated during the observation period)
  - Period: 30 days of June 2006

Experiments on MSN

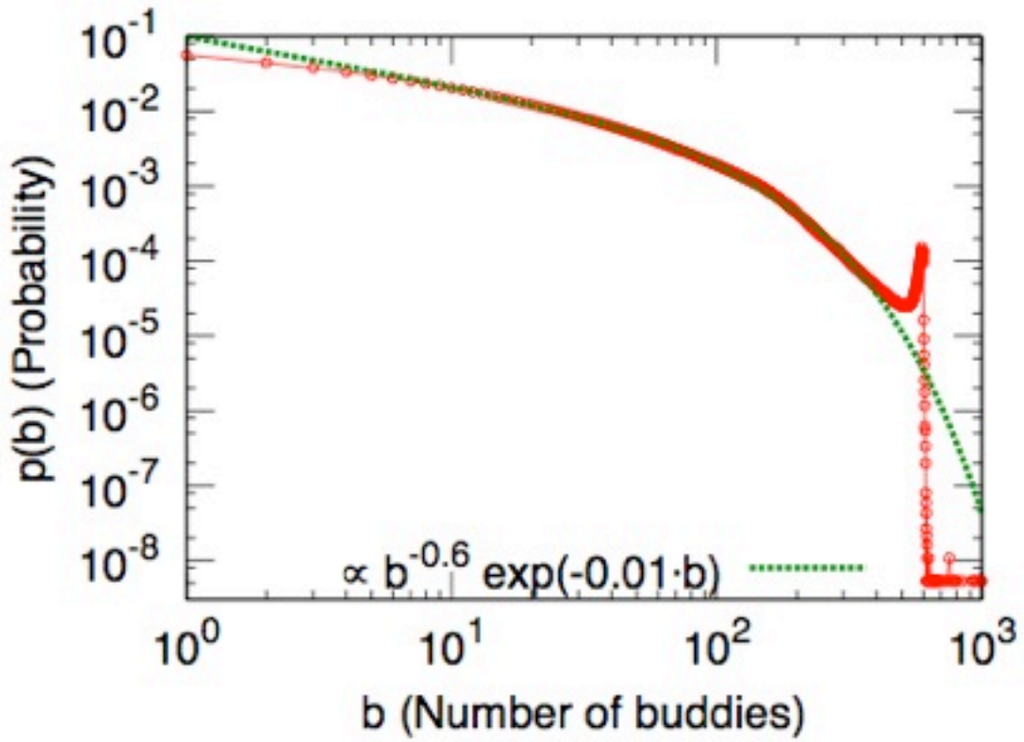
[Jury Leskovec, 2008]

Heavy tailed but not power-law distribution



(a) Communication

Degree distribution of communication network (num. of people with whom a person communicates)



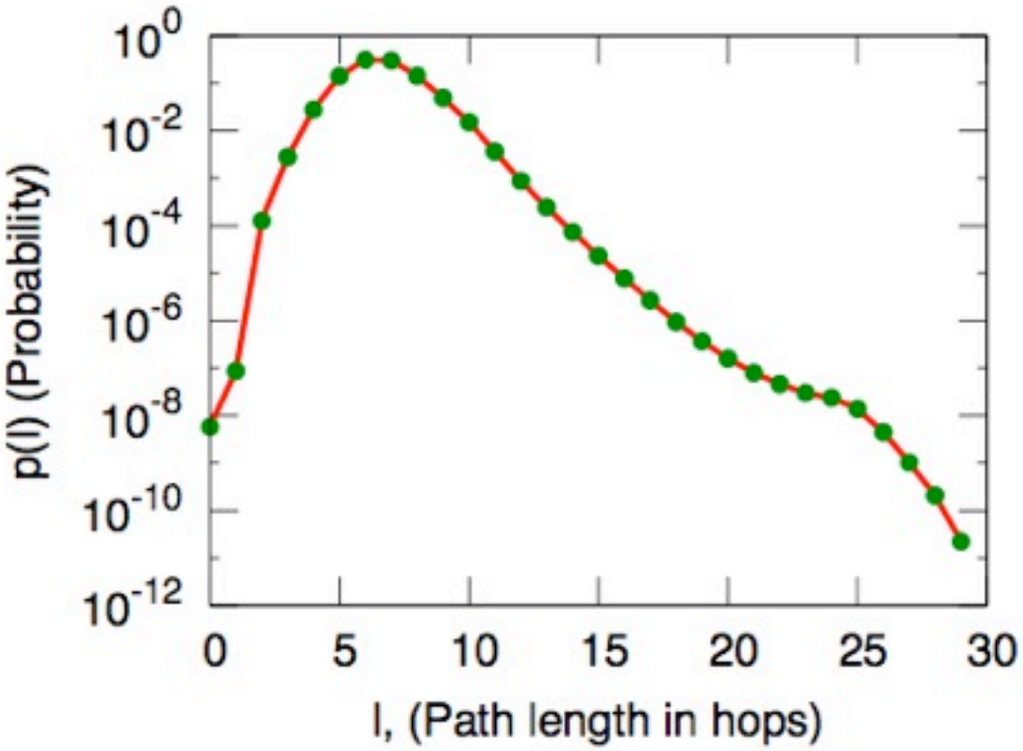
(b) Buddies

Degree distribution of the buddy network (length of the contact list)

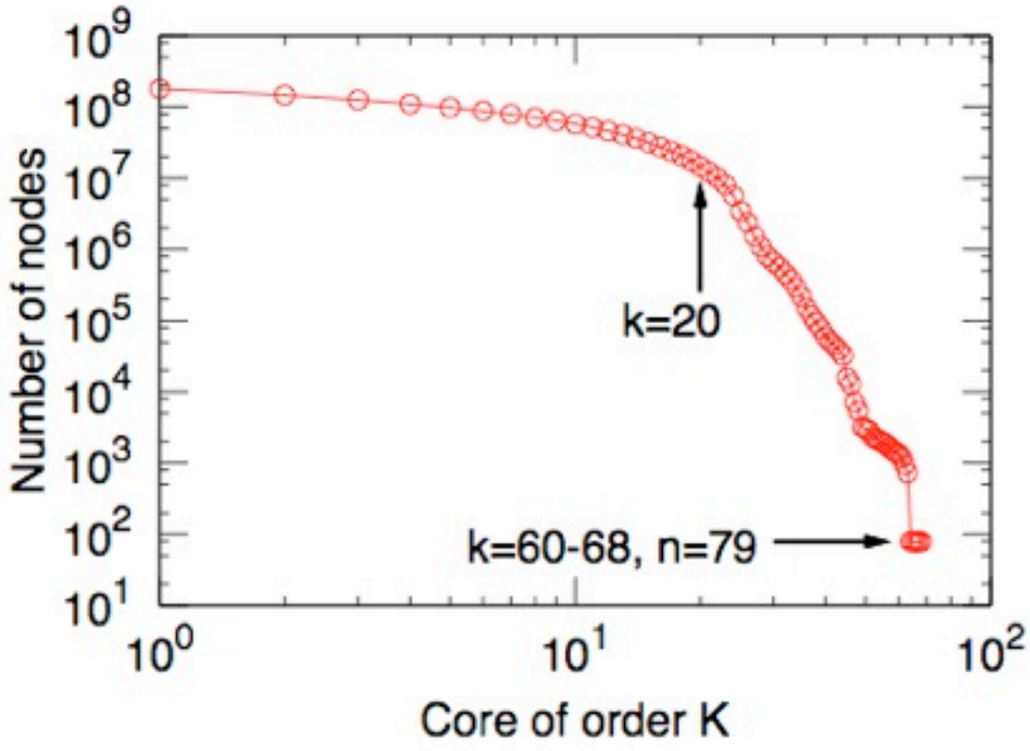
Experiments on MSN

[Jury Leskovec, 2008]

Average shortest path has length 6.6



Distribution over the shortest path lengths



Distribution of sizes of cores of oder k

Experiments on Mobile Network

[Zheng-Bin Dong, 2009]

- 2,590,361 phone users were extracted from one month call logs of a city in China
  - The mobile social network with 2,590,361 nodes is not a connected graph, which has 56,601 connected subgraphs
  - The largest connected subgraph has 2,532,298 nodes – the percentage is 97.76%
  - The analysis is focus on the largest connected subgraph

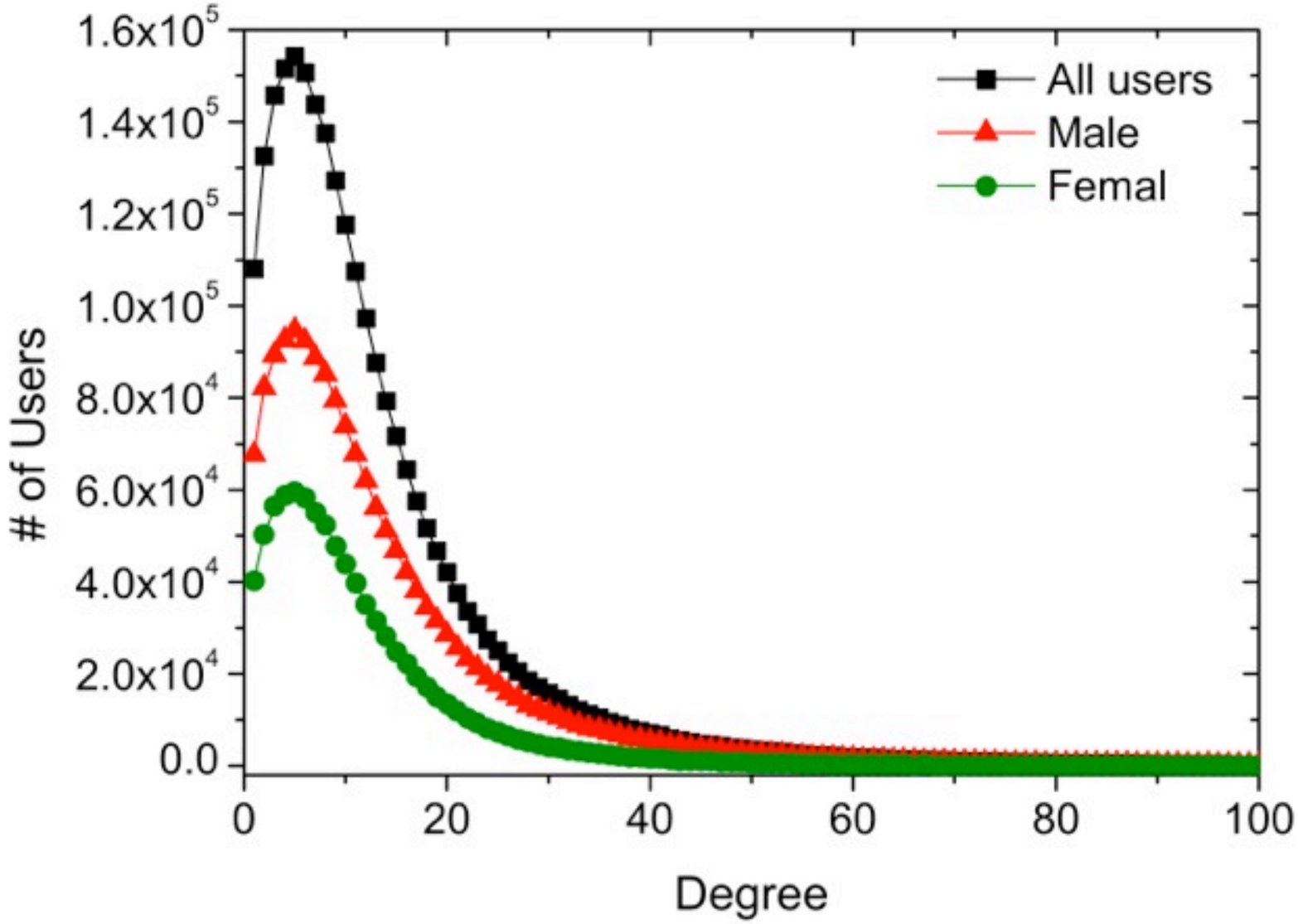
Six Degree Facebook

[Wiki]

- It calculates the degrees of separation between different people
- It has 4.5 million users as of April 7, 2008
- The average separation for all users of the application is 5.73 degrees, whereas the maximum degree of separation is 12

Experiments on Mobile Network

[Zheng-Bin Dong, 2009]

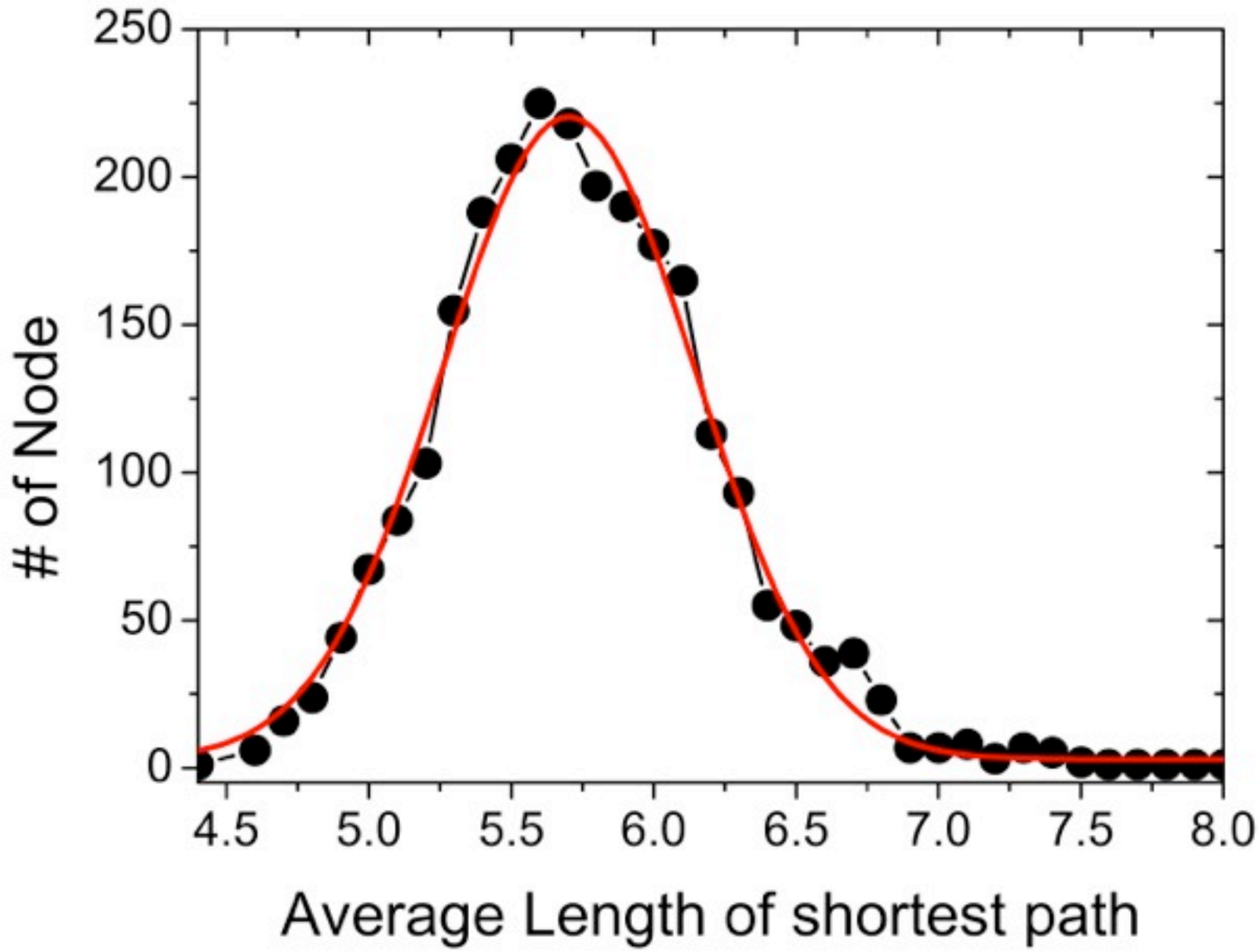


Degree distribution

Experiments on Mobile Network

[Zheng-Bin Dong, 2009]

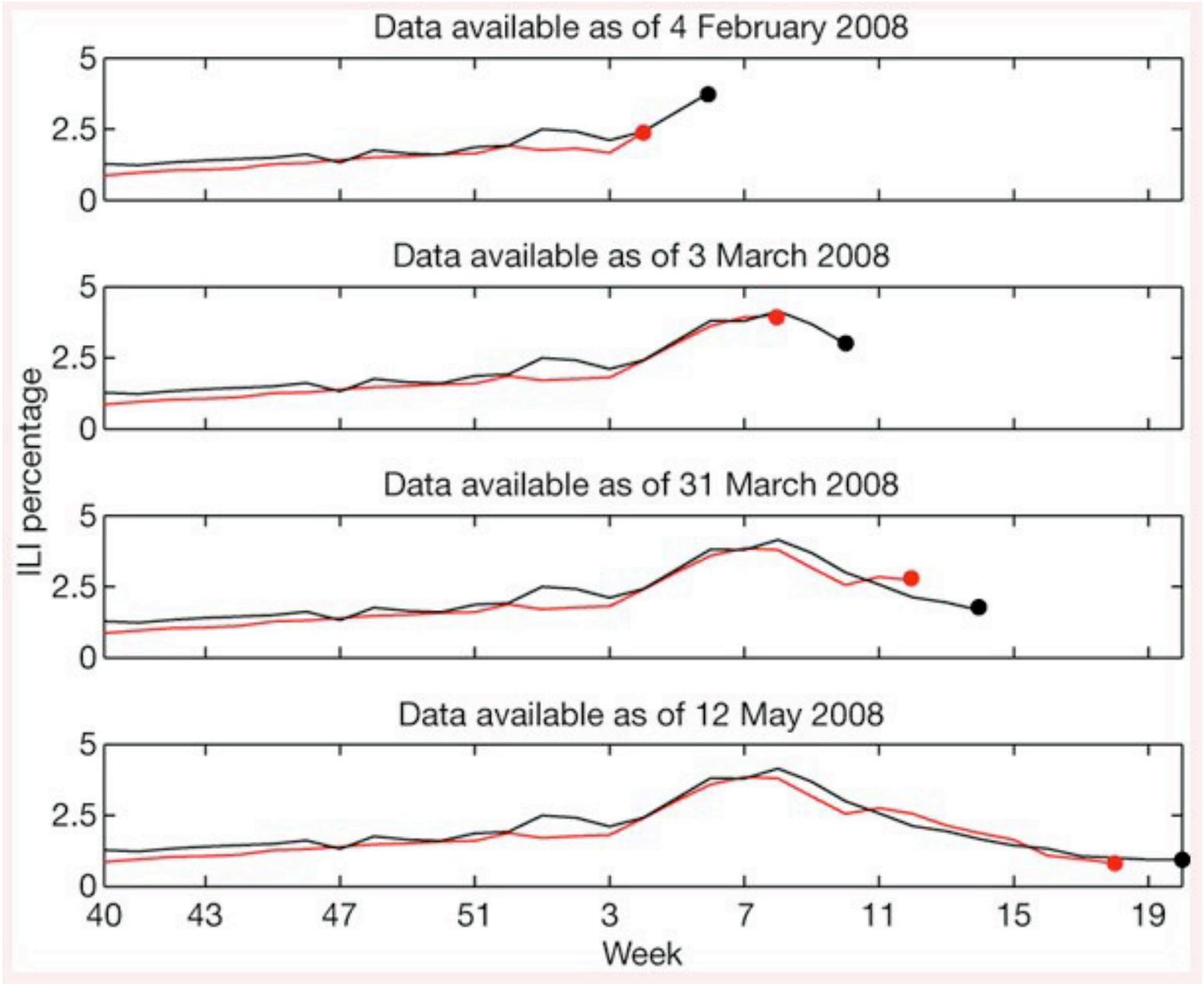
Overall average shortest path is 5.75



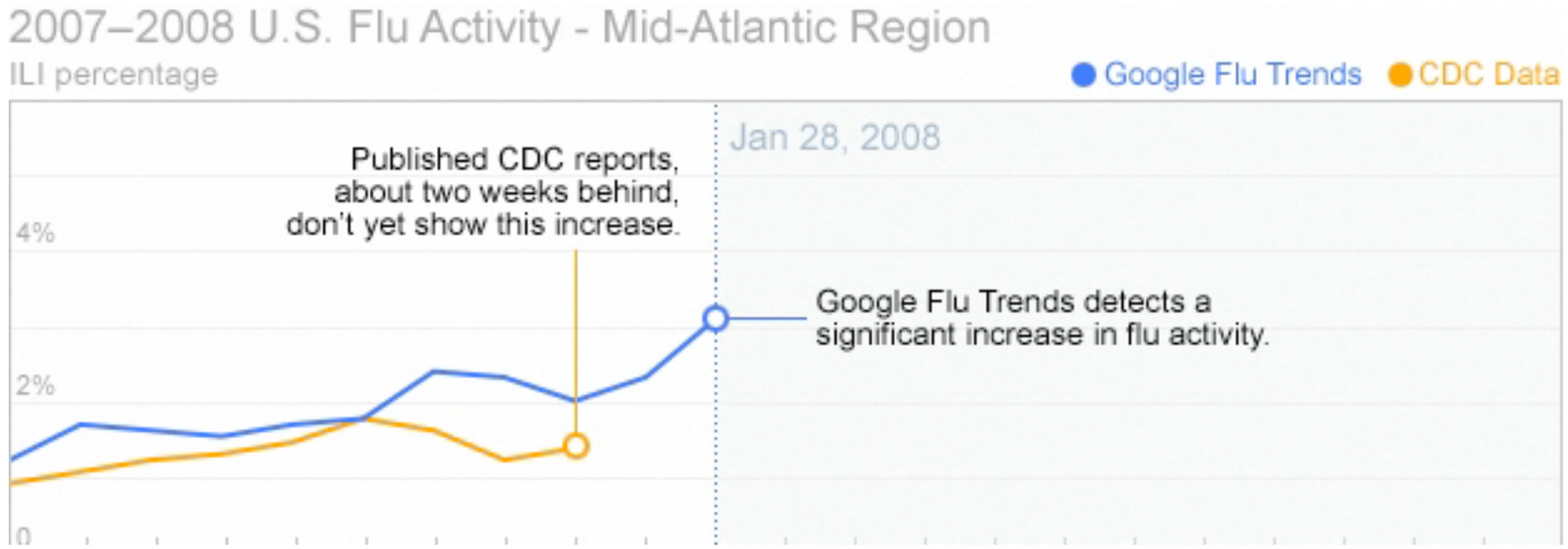
Characteristics of Social Networks

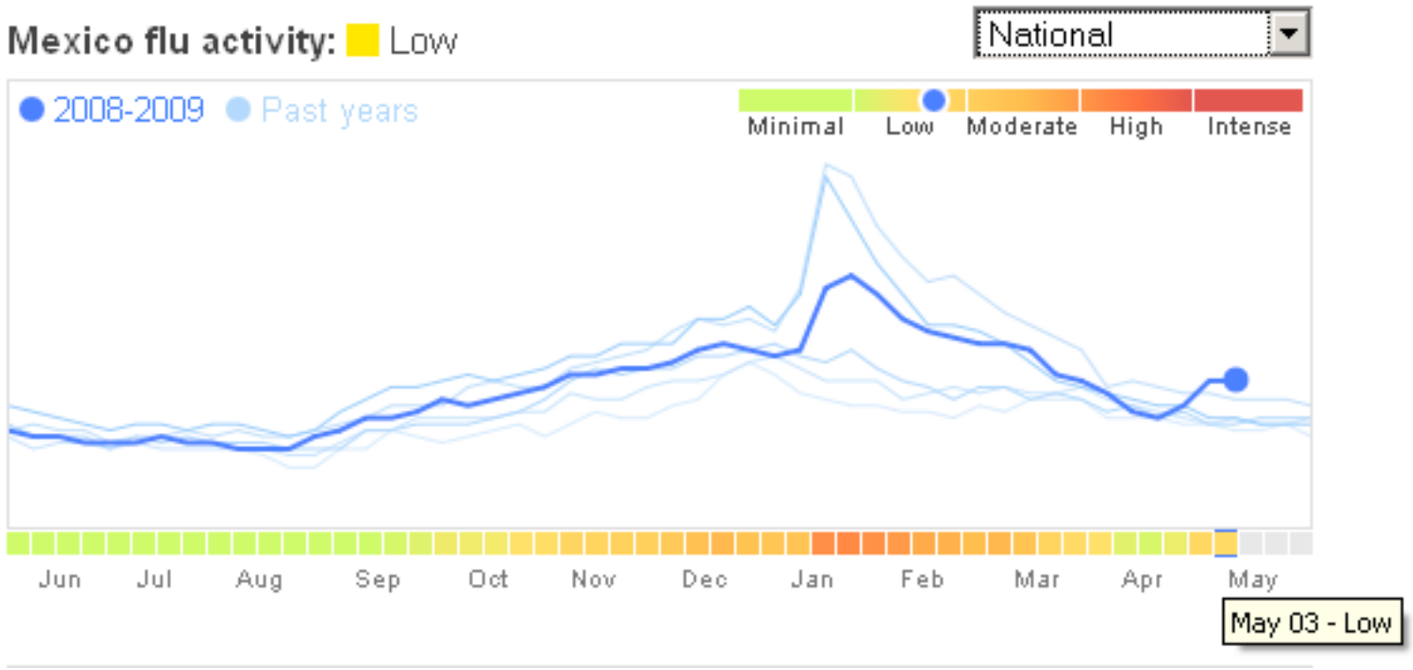
		Power law distribution		Average shortest path	
		In-degree	Out-degree	Directed	Undirected
Co-authorship network		2~3		6~16	
Email network		5.9	5.2	6	
World Wide Web		2.1	2.7	16	6
Blog	Blog network	1.7	N/A	N/A	
	Post network	2.15	2.95		
Instant-messaging network		Heavy tailed but not power law		6.6	
Facebook		N/A		5.73	
Mobile network		N/A		5.75	

Google' Plague Prediction System



Google's Plague Prediction System

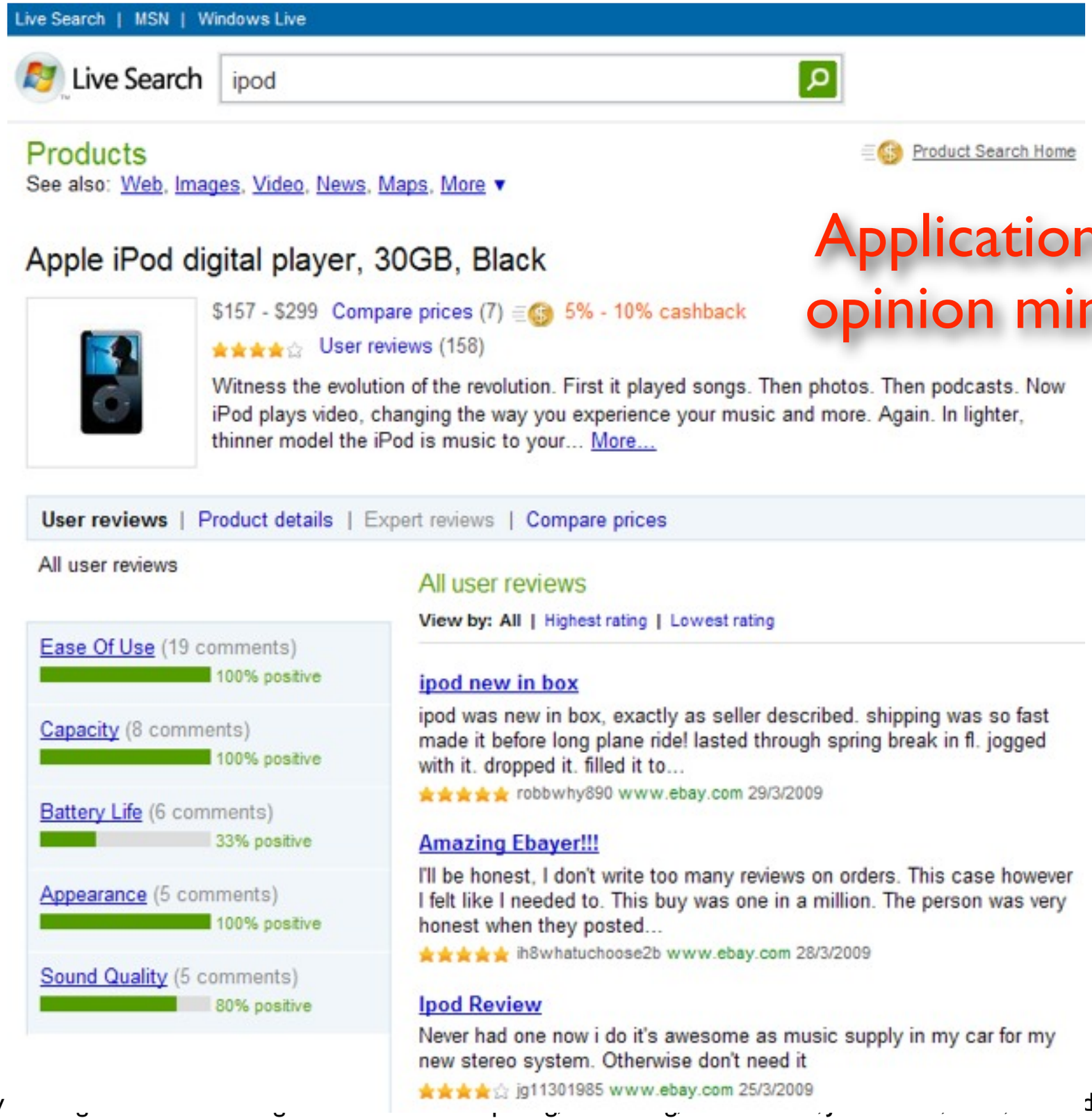




Google's search trends can predict flu epidemic decline

The New Paradigm Shift: The Emergence of Social Computing, Irwin King, IWSSIP2009, June 18-20, 2009, Chalkida, Greece

Live Search



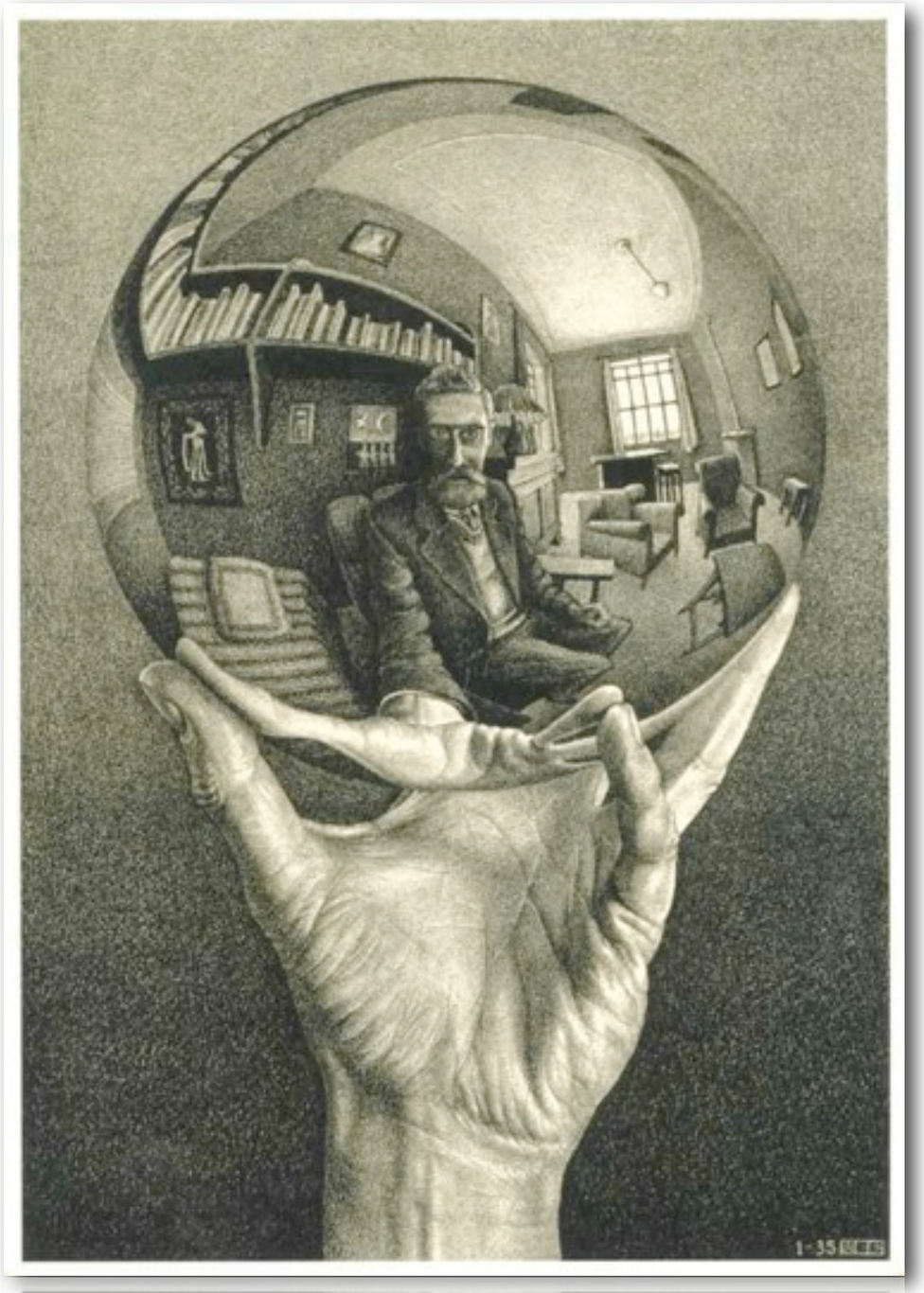
Application of opinion mining

Summary

- The new Social Computing paradigm
- Dealing with social behavior
- Partial and incomplete information
- Spatial and temporal data
- Collective wisdom

What's on the Horizon

- Theory and models
- Scalability and algorithmic issues
- Security and privacy issues
- CLOUD (broadband + wireless)
- Web Services
- Monetization of Social Interactions



On-Going Research

Machine Learning

- Direct Zero-norm Optimization for Feature Selection (ICDM'08)
- Semi-supervised Learning from General Unlabeled Data (ICDM'08)
- Learning with Consistency between Inductive Functions and Kernels (NIPS'08)
- An Extended Level Method for Efficient Multiple Kernel Learning (NIPS'08)
- Semi-supervised Text Categorization by Active Search (CIKM'08)
- Transductive Support Vector Machine (NIPS'07)
- Global and local learning (ICML'04, JMLR'04)

Web Intelligence

- Effective Latent Space Graph-based Re-ranking Model with Global Consistency (WSDM'09)
- Formal Models for Expert Finding on DBLP Bibliography Data (ICDM'08)
- Learning Latent Semantic Relations from Query Logs for Query Suggestion (CIKM'08)
- RATE: a Review of Reviewers in a Manuscript Review Process (WI'08)
- MatchSim: link-based web page similarity measurements (WI'07)
- Diffusion rank: Ranking web pages based on heat diffusion equations (SIGIR'07)
- Web text classification (WWW'07)

Collaborative Filtering

- Recommender system: accurate recommendation based on sparse matrix (SIGIR'07)
- SoRec: Social Recommendation Using Probabilistic Matrix Factorization (CIKM'08)

Human Computation

- An Analytical Study of Puzzle Selection Strategies for the ESP Game (WI'08)
- An Analytical Approach to Optimizing The Utility of ESP Games (WI'08)

Acknowledgments

- Prof. Michael R. Lyu
- Prof. Jimmy Lee
- Dr. Zenglin Xu
- Dr. Zhirong Yang
- Thomas Chan (M.Phil)
- Hongbo Deng (Ph.D.)
- Zhenjiang Lin (Ph.D.)
- Hao Ma (Ph.D.)
- Haiqin Yang (Ph.D.)
- Xin Xin (Ph.D.)
- Chao Zhou (Ph.D.)

Workshop on Social Computing in Education 2009

Home

New since last time: 1 file

http://groups.google.com/group/WSCE2009

Call for Papers

Workshop on Social Computing in Education (WSCE2009)
in conjunction with SocialComp-09, August 29-31, 2009, Vancouver, Canada

Welcome to the workshop on Social Computing in Education (SCE2009). The workshop is held in conjunction with the SocialComp-09, Vancouver, Canada from August 29-31, 2009.

With the advent of Web 2.0 and related technologies, Social Computing has become a new paradigm in ways we communicate, learn, and educate. Social platforms such as wikis, blogs, twitters, forums, groups, podcasts, mashups, virtual worlds, and sites for social networking, recommender systems, social bookmarking, social news, knowledge sharing, etc. are generating novel ways we acquire, access, manipulate, process, retrieve, present, and visualize information in the teaching and learning space. The social media for education has become dynamic, ubiquitous, distributed, real-time, collaborative, bottom-up, many-to-many, value-based, and personalized. This workshop solicits contributions on using Social Computing and related technologies for education, the emerging applications of Web 2.0 as an educational platform, as well as privacy, risk, security, and policy issues associated in Social Computing for Education 2.0.

King · Baeza-Yates (Eds.)

Weaving Services and People on the World Wide Web

Ever since its inception, the Web has changed the landscape of human experiences on how we interact with one another and data through service infrastructures via various computing devices. This interweaving environment is now becoming ever more embedded into devices and systems that integrate seamlessly on how we live, both in our working or leisure time.

For this volume, King and Baeza-Yates selected some pioneering and cutting-edge research work that is pointing to the future of the Web. Based on the Workshop Track of the 17th International World Wide Web Conference (WWW2008) in Beijing, they selected the top contributions and asked the authors to resubmit their work with a minimum of one third of additional material from their original workshop manuscripts to be considered for this volume. After a second-round of reviews and selection, 16 contributions were finally accepted.

The work within this volume represents the tip of an iceberg of the many exciting advancements on the WWW. It covers topics like semantic web services, location-based and mobile applications, personalized and context-dependent user interfaces, social networks, and folksonomies. The presentations aim at researchers in academia and industry by showcasing latest research findings. Overall they deliver an excellent picture of the current state-of-the-art, and will also serve as the basis for ongoing research discussions and point to new directions.

ISBN 978-3-642-00569-5

springer.com

King · Baeza-Yates (Eds.)

Weaving Services and People on the World Wide Web

Irwin King
Ricardo Baeza-Yates (Eds.)

Weaving Services and People on the World Wide Web

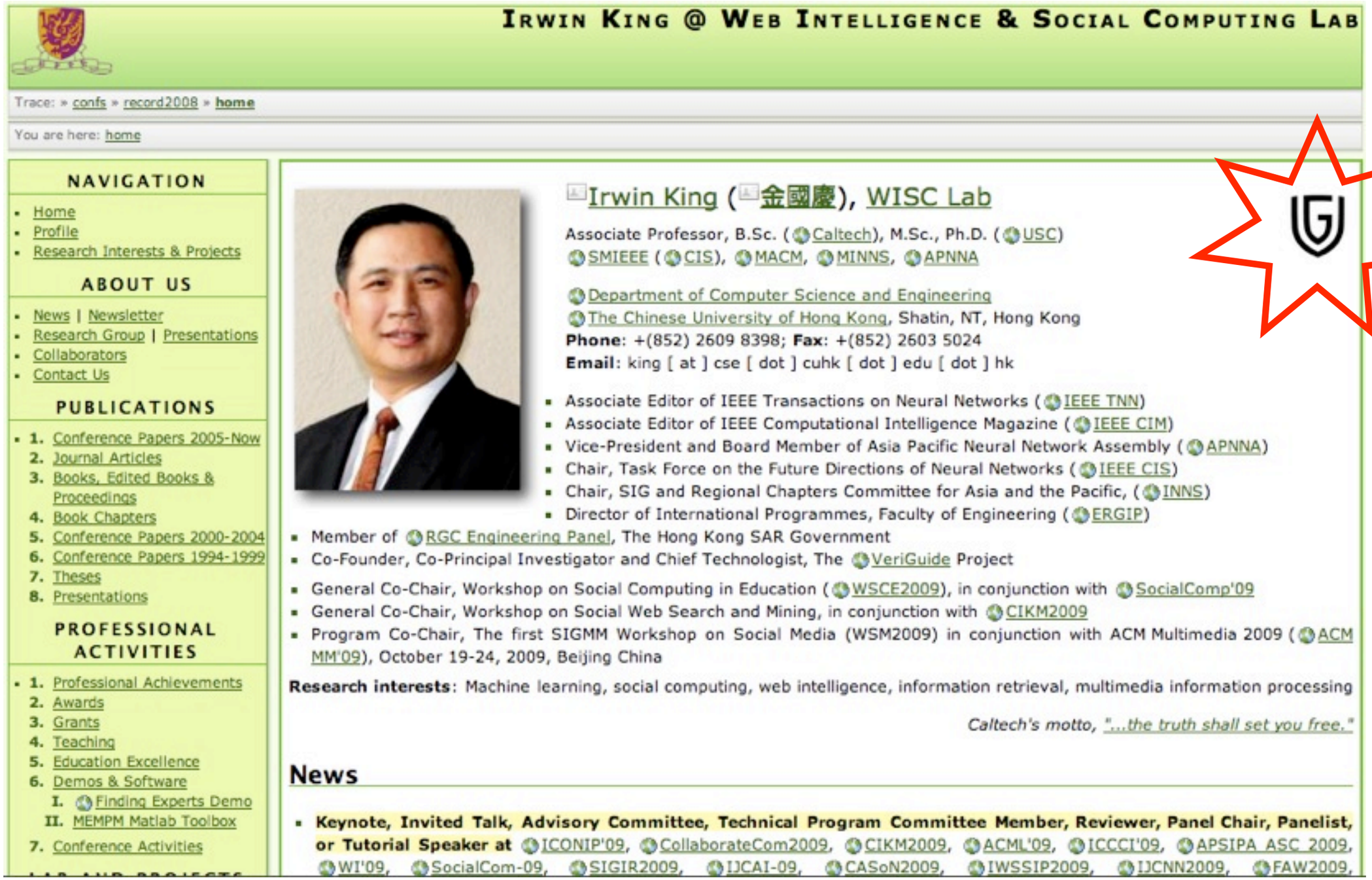
Springer

The New Paradigm Shift: The Emergence of Social Computing, Irwin King, IWSSIP2009, June 18-20, 2009, Chalkida, Greece

VeriGuide

- Similarity text detection system
- Developed at CUHK
- Promote and uphold academic honesty, integrity, and quality
- Support English, Traditional and Simplified Chinese
- Handle .doc, .txt, .pdf, .html, etc. file formats
- Generate detailed originality report including readability

VeriGuide Free Trial



http://www.cse.cuhk.edu.hk/~king

IWSSIP2009

References

- A.L. Barabasi, H. Jeonga , Z. Neda, E. Ravasz, A. Schubert, and T. Vicsek, Evolution of the social network of scientific collaborations, 2002
- Diane Lambert. Communication Flows: Analysis of an Email Network
- Chunguang Li and Guanrong Chen. Network connection strengths: Another power-law?
- L. von Ahn and L. Dabbish. Labeling images with a computer game. 2004
- Andrei Broder, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, and Janet Wiener. Graph structure in the web. 2000
- Jure Leskovec, Mary McGlohon, Christos Faloutsos, Natalie Glance, and Matthew Hurst. Cascading Behavior in Large Blog Graph. 2007
- Jure Leskovec and Eric Horvitz. Planetary-Scale Views on a Large Instant-Messaging Network. 2008
- Zheng-Bin Dong, Guo-Jie Song, Kun-Qing Xie, and Jing-Yao Wang. An Experimental Study of Large-Scale Mobile Social Network. 2009