

ICT & Social Epistemology

L1. Filtering and Ranking Information

Management of Organisations
ESC Dijon Bourgogne
1-5 March 2010

Dario Taraborelli
Centre for Research in Social Simulation
University of Surrey

Living in an attention economy

“In an information-rich world, the wealth of information means a dearth of something else: a scarcity of whatever it is that information consumes. **What information consumes is** rather obvious: it consumes **the attention of its recipients**. Hence a wealth of information creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it”

Herbert A. Simon (1971)

Designing Organizations for an Information-Rich World

Information pollution and information overload

“The **Internet is the worst polluter** of all. Spam isn't even pollution, it's attention theft. But even legitimate email is typically copied to more people than necessary and contaminated by excess verbiage and endless reply loops. Studies of content usability typically find that removing half of a website's words will double the amount of information that users actually get.”

Jakob Nielsen (2003)
Information Pollution

Information snacking

“The easier it is to find places with good information, the **less time** users will spend visiting any individual website. This is one of many conclusions that follow from analyzing how people optimize their behavior in online information systems.”

This is the reason why **effective ranking solutions makes people leave websites faster**

Jakob Nielsen (2003)
Information Foraging

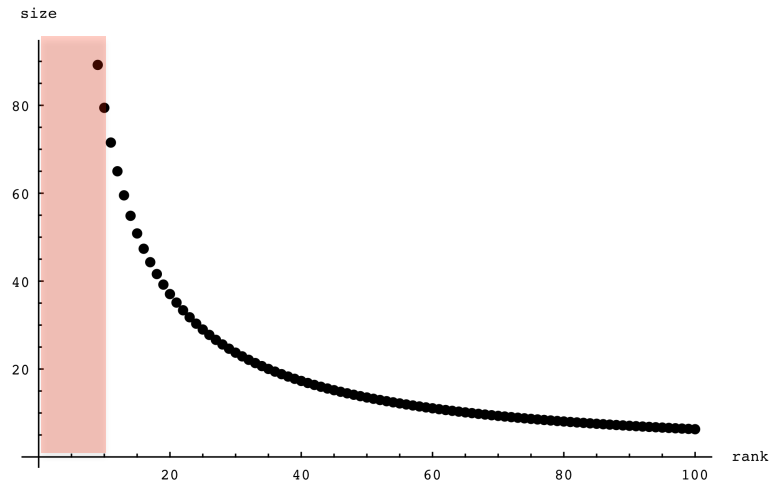
The role of ranking technology

flickr 3,396 uploads in [the last minute](#)

- Nowadays information technology (in particular reputation and ranking technology) controls the economy of attention.
- The function of ranking technology:
 1. Control the largest possible share of user attention
 2. Optimise information snacking
 3. Preserve (or reduce) diversity to suit user needs

Power laws in the Web

Typical long-tail distribution of many variables in the Web (number of incoming links, number of comments, number of bookmarks, number of visits, rank of search results visited)



Ranking technology to:

- yield the most authoritative items
- preserve diversity

Ranking information

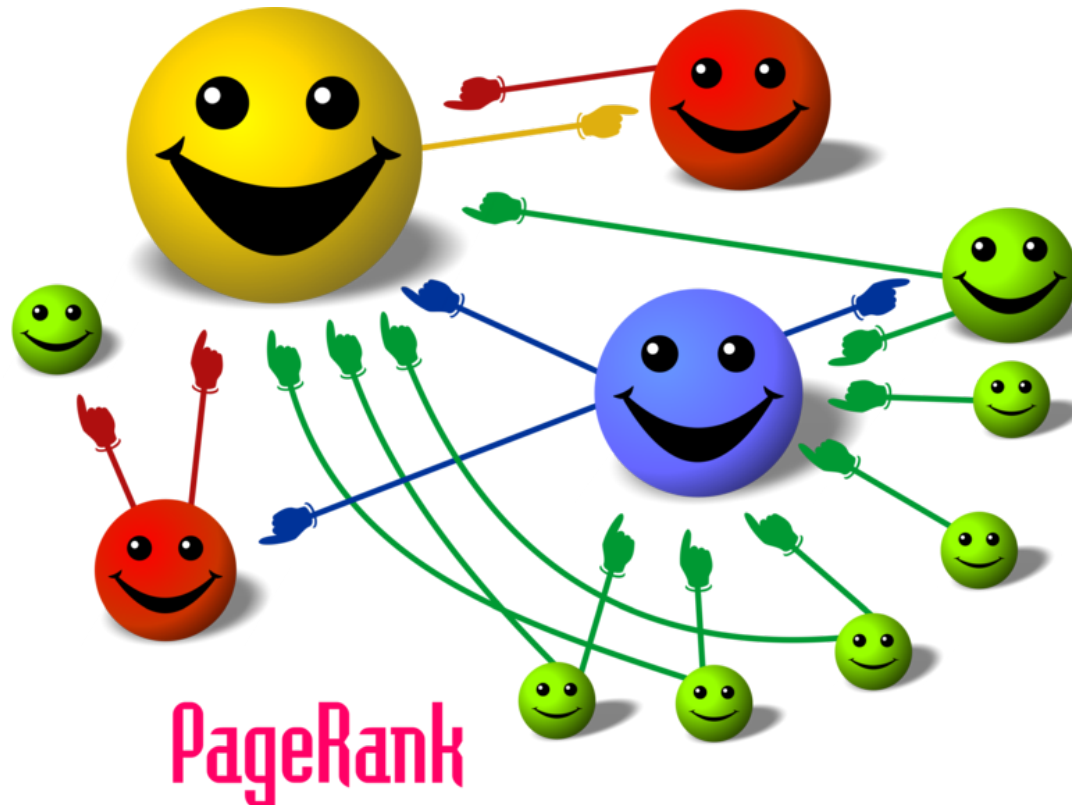
A short overview of ranking & filtering solutions:

- First-generation search engines
- Google's PageRank
- The Web 2.0: collective attention
- The Web 2.0: collaborative ranking
- Social and personalised search
- Online reputation systems
- Distributed ranking systems

First-generation search engines

VIAGRA PHENTERMINE LORAZEPAM CIALIS XANAX TRAMADOL
HYDROCHLORIDE CHEAP VIAGRA PHENTERMINE LORAZEPAM CIALIS
XANAX TRAMADOL HYDROCHLORIDE CHEAP VIAGRA PHENTERMINE
LORAZEPAM CIALIS XANAX TRAMADOL HYDROCHLORIDE CHEAP
**Failure of early ranking solutions
based on keyword frequency**
VIAGRA PHENTERMINE LORAZEPAM CIALIS XANAX TRAMADOL
HYDROCHLORIDE CHEAP VIAGRA PHENTERMINE LORAZEPAM CIALIS
XANAX TRAMADOL HYDROCHLORIDE CHEAP VIAGRA PHENTERMINE
LORAZEPAM CIALIS XANAX TRAMADOL HYDROCHLORIDE CHEAP
VIAGRA PHENTERMINE LORAZEPAM CIALIS XANAX TRAMADOL
HYDROCHLORIDE CHEAP PHENTERMINE LORAZEPAM CIALIS XANAX
TRAMADOL HYDROCHLORIDE CHEAP PHENTERMINE LORAZEPAM CIALIS
XANAX TRAMADOL HYDROCHLORIDE CHEAP PHENTERMINE LORAZEPAM

Google PageRank

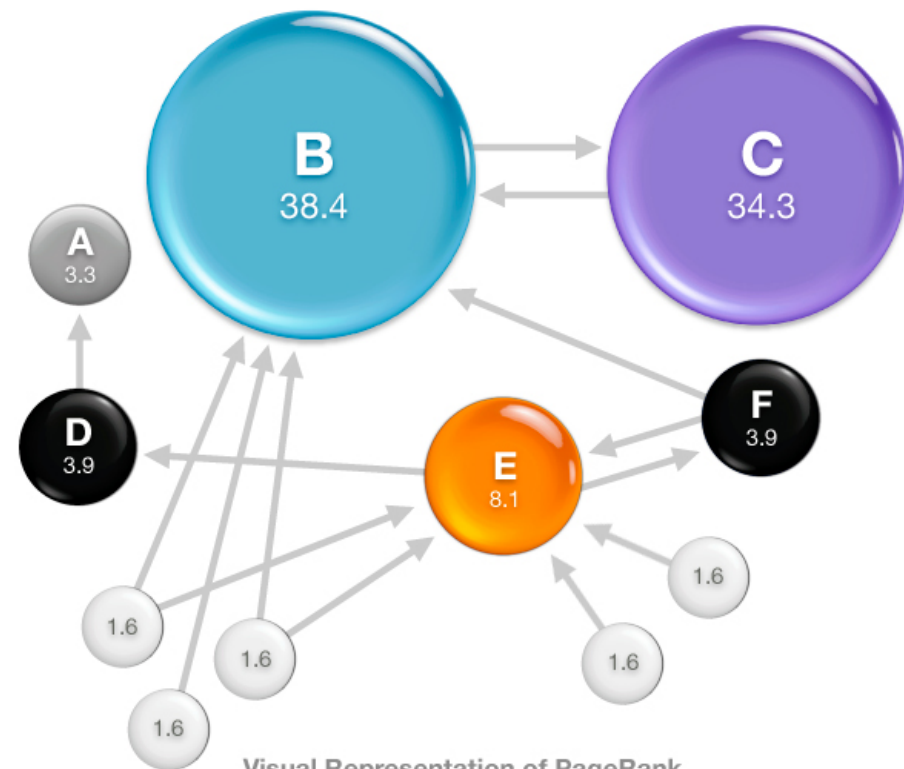


Sergei Brin and Larry Page (1998)
*The anatomy of a large-scale hypertextual
Web search engine*

Google PageRank

A website's **authority** is determined by (the weighted sum of) all known websites that link to it (multiplied by their respective authority)

Idea inspired by the study of **citations: linking is a form of endorsement**



Google PageRank

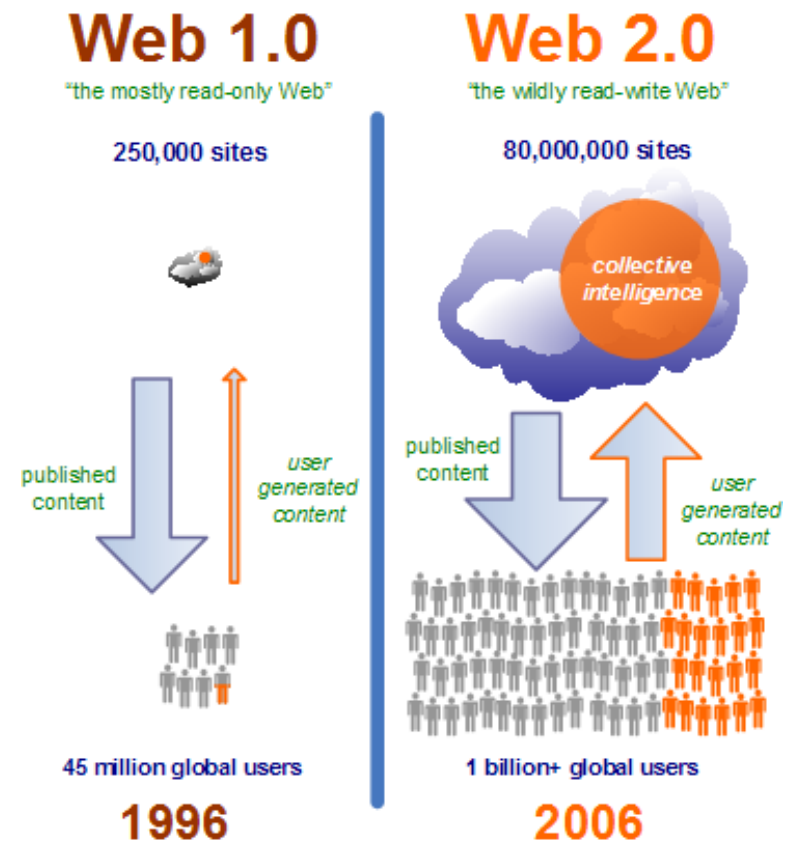
- **Robustness** of the system: to be able to game PageRank one needs to either:
 - have a high reputation
 - be able to control a very large number of lower reputation websites (*Google Bombs*)
- **Search engine optimisation** industry

The Web 2.0 turn



The Web 2.0 turn


- From the **read-only Web** to the massively **read-write Web**
- Radically **distributed** information selection
- Crucial role of **aggregating algorithms**






The Web 2.0 turn




- Collective attention
- Distributed information selection

News, Images, Videos Most Recent Top in 24 Hr 7 Days 30 Days 365 Days


210 diggs [Tsunami Warnings After Massive 8.5 Earthquake Strikes Chile](#) 
dailymail.co.uk — Witnesses reported collapsed buildings and tremors lasting up to 90 seconds after the quake, which struck at around 6.30am GMT..... (Submitted by FISHERMAN)




 19 Comments  Share  Bury Made popular **14 min ago**

128 diggs [12 Secrets Of Effective Business Communication](#)
noupe.com — You could be an excellent designer, but if you're unable to promote your services and communicate effectively with clients and colleagues, your potential is limited. (Submitted by ImALiar)

 6 Comments  Share  Bury Made popular **24 min ago**

Sponsored by Toyota

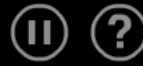
420 diggs  [What to do if your accelerator pedal sticks](#)
toyota.com - Learn how to bring your Toyota vehicle to a safe stop if you experience a sticking accelerator pedal.

  Share  Bury

stack

real-time activity

zoom: 100



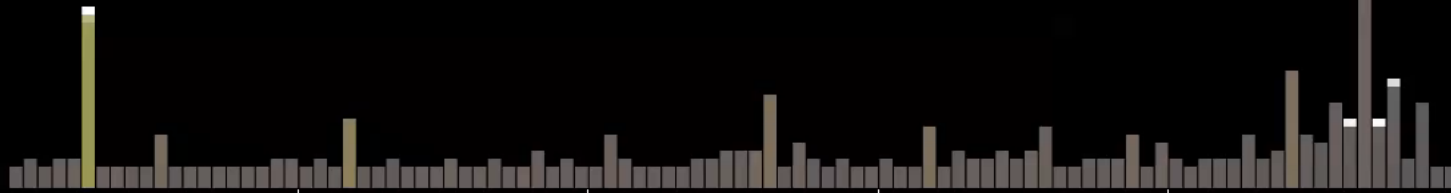
nalen33

Hoopsterm91

BZKyle
jeremyosborne81

BZKy

5:43:17
pm
start



How NOT to Hide Cocaine in Your Car (PICS)
The Global e-waste Crisis (Slideshow)
Where The Fuck Are My Cookies? [Pic]

Sony, LG, Samsung, Hitachi, Toshiba accused of price fixing
Easter Eggs for Hitler
Couple Receives 30 Pounds of Pot from FedEx
A massive magnitude-8.3 earthquake struck Chile
"Vampire of Venice" Unmasked: Plague Victim & Witch?
Fallout Artist Reveals Original Childkiller Vault-Boy Image
Toxic towns: People of Mossville 'are like an experiment'

Guinea Pigs are Awesome [15 Pics]
Thank you Digg - The Oatmeal
Easter Eggs for Hitler

5:51:10
pm
now

diggings 0 2500+

The Web 2.0 turn

- Collaborative ranking

flickr

Explore / Interestingness / January 2010 / 26th

« [25 January 2010](#)

[26 January 2010](#)



The Scout

Uploaded on 26 January 2010



By [Rian Flynn](#)
His [photostream](#), or [profile](#).

 [the](#), [scout](#), [laurenmcknight](#), [esnipecom](#) ...

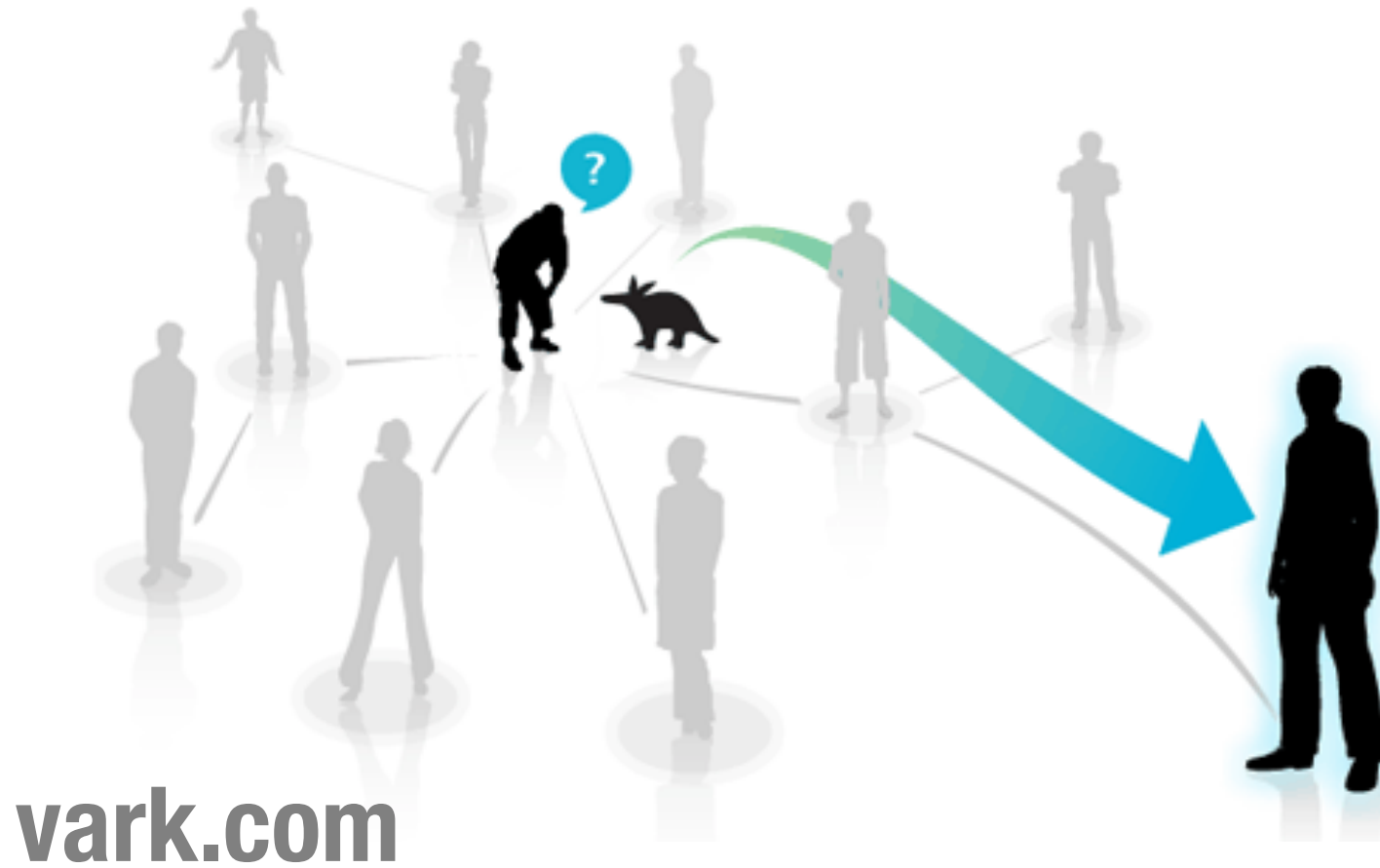


rainy bench monday

Social and personalised search

- **Personalised search**
 - use one's own **search history** to improve ranking
- **Social search**
 - use one's **social network** to improve the relevance of results

Social search



Online reputation systems

Electronic markets (*eBay*)

Feedback Profile



wheels_4_babies (2005 ★) me

Positive Feedback (last 12 months): 99.6%
[\[How is Feedback Percentage calculated?\]](#)

Member since: 25-Jul-07 in Germany

Recent Feedback Ratings (last 12 months)

	1 month	6 months	12 months
Positive	140	660	1346
Neutral	1	6	11
Negative	1	4	5

Detailed Seller Ratings (last 12 months)

Criteria	Average rating	Number of ratings
Item as described		1058
Communication		1061
Dispatch time		969
Postage and packaging charges		1054

Online reputation systems

Collaborative problem solving (*StackOverflow*)



RedGlyph



2,715

reputation

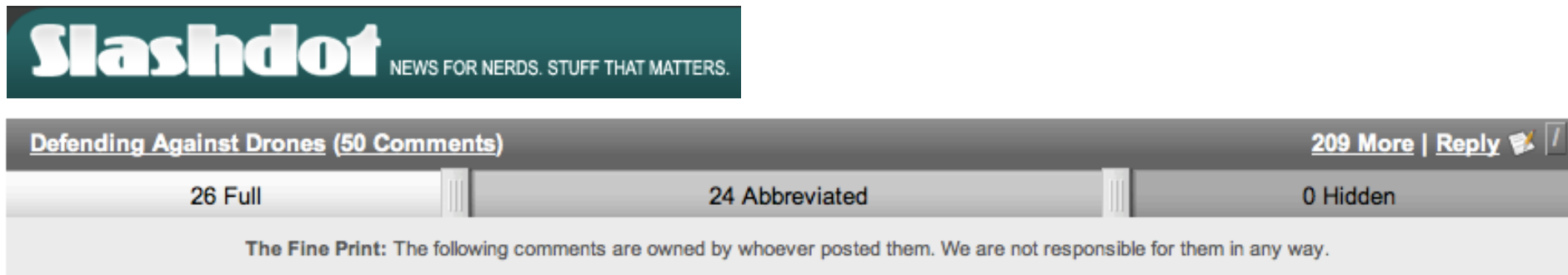
107 views

Registered User

name	RedGlyph
member for	4 months
seen	9 hours ago
website	
location	
age	

Online reputation systems

Collaborative news selection (*Slashdot*)



Slashdot NEWS FOR NERDS. STUFF THAT MATTERS.

Defending Against Drones (50 Comments) 209 More | Reply

26 Full | 24 Abbreviated | 0 Hidden

The Fine Print: The following comments are owned by whoever posted them. We are not responsible for them in any way.

➤ [Hey...](#) (Score:2, Funny) by Anonymous Coward what about bullets? They have been in the market for quite long already.

↳ [Re:Hey... bullets!](#) (Score:3, Informative) by [thms \(1339227\)](#) Not like they used to. [Air burst rounds](#) [wikipedia.org] will

Arm your citizens... (Score:3, Informative)

by [saleenS281 \(859657\)](#) on Saturday February 27, @09:26AM (#31296502) [Homepage](#)

It would seem to me if every citizen knew how to properly shoot a rifle, odds are pretty good one of those things could be knocked out of the sky with a barrett. It would cost all of us a heck of a lot less money too.

In fact... this is exactly the sort of thing the 2nd amendment was written for. "The people" defending themselves from attack.

Distributed ranking systems

Issues in the design of distributed ranking systems:

1. radically distributed evaluation
2. bottom-up definition of authority
3. control exerted through aggregating algorithms
4. risks of gaming and free-riding

Questions

- How to build an **effective** distributed ranking system on the basis of available information?
- How to make a distributed ranking system “**interesting**” (preserving diversity)?
- How to make a distributed ranking system **game-resistant**?