



# Web 2.0

Professor Fabio Ciravegna  
 Department of Computer Science  
 University of Sheffield  
[fabio@dcs.shef.ac.uk](mailto:fabio@dcs.shef.ac.uk)

COM3504/6504/6505  
 “The Intelligent Web”  
<http://www.dcs.shef.ac.uk/~cirav/teaching/>

# Paper faces legal threat over picture of footballer

Anushka Asthana, Frances Gibb and Alexi Mostrouss

May 23 2011 12:01AM

The Attorney-General is considering contempt proceedings against a Sunday newspaper that has effectively identified a footballer behind a controversial privacy case. The footballer has already been named in tens of thousands of Twitter messages — up to one tweet every five seconds yesterday contained his name — and his lawyers have begun action against the micro-blogging site in their hunt for the Twitter user who first disclosed his identity. But yesterday the player's legal fight escalated with the publication of the photograph by the Scottish Sunday Herald — with the player's eyes blacked out and “censored” written across them. The move by the Attorney-General to take action against a Scottish newspaper beyond the remit of the English courts adds a further twist to the legal row over injunctions. Dominic Grieve, QC is also expected to consider contempt proceedings against a journalist...

, University of Sheffield





Two university spin-outs working on Web 2.0, mobile phone apps, and web technologies are offering a couple of internships each for the summer

3<sup>rd</sup> or 4<sup>th</sup> year  
starting mid/end of May

The Floow Ltd <http://www.thefloow.com/>

K-Now Ltd <http://www.k-now.co.uk/>

*All the personnel is composed of former students. They have hired a dozen of them so far (including one in this room!)*

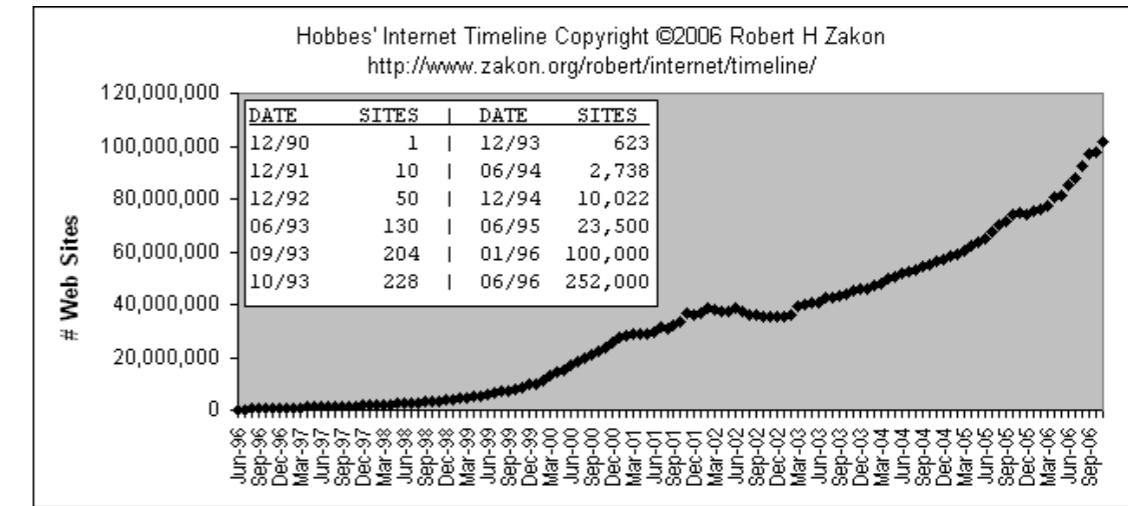
Based on:

- A customer (you)
- A program to access the service (the browser)
  - Separation from the rest of the desktop
- A proprietary program and associated knowledge (e.g. an e-commerce Web site)



The  
University  
Of  
Sheffield.

# Web 2.0



1991

1995

2000

2005

2009

## Issues

- Harnessing collective Intelligence
- Trusting users as co-developers
- Creating, supporting and free communities



Jan 2001  
Wikipedia

April 2003  
Apple's iTunes: bound to become the largest store in  
the world for digital music



1999



2002



2004



Source: BBC



2005

# Web 2.0 in 5 minutes

- The video "Web 2.0 ... The Machine is Us/ing Us."
  - Was originally released on YouTube on January 31st 2007
  - Quickly became the most popular video in the blogosphere and the #1 featured YouTube video on February 7th 2007.



# Web 2.0: Trends

**What Is Web 2.0**

Design Patterns and Business Models for the New Generation of Software

by Tim O'Reilly  
09/30/2005

- From connecting documents to connecting people
- User trends
  - Harnessing collective intelligence
  - Trusting users as co-developers
  - Leveraging the long tail through customer self-service

<http://oreilly.com/web2/archive/what-is-web-20.html>

The diagram illustrates the evolution of social media. It starts with a screenshot of a Facebook profile page on the left, showing a timeline of posts and interactions. A pink arrow points from this Facebook page to a central Twitter logo. Below the Twitter logo is a screenshot of a Facebook group page for 'shefcompsci'. A green arrow points from the Twitter logo down to the shefcompsci group page. The shefcompsci page displays basic group info, recent news, members, and related groups. At the bottom of the page, there is a URL: <http://www.oreillynet.com/pwo/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>.

# Web 2.0: Trends

**What Is Web 2.0**  
Design Patterns and Business Models for the Next  
Generation of Software

by Tim O'Reilly  
09/30/2005

- Products

- Services, not packaged software, with cost-effective scalability
- Control over unique, hard-to-recreate data sources that get richer as more people use them
- Software above the level of a single device
- Lightweight user interfaces, development models, AND business models

# Harnessing collective intelligence

- Hyperlinking is the foundation of the web.
  - As users add new content, and new sites, it is bound in to the structure of the web by other users discovering the content and linking to it.
    - Much as synapses form in the brain, with associations becoming stronger through repetition or intensity, the web of connections grows organically as an output of the collective activity of all web users.
- Google's breakthrough in search was PageRank
  - A method of using the link structure of the web rather than just the characteristics of documents to provide better search results.
  - What is popular is important
    - Many hyperlinks reaching the page
    - Important hyperlinks reaching the page

**What Is Web 2.0**  
Design Patterns and Business Models for the Next Generation of Software

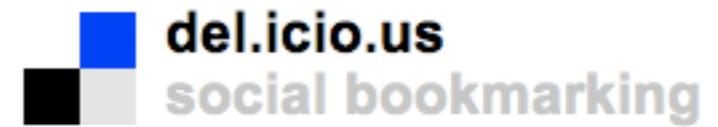
by Tim O'Reilly  
09/30/2005

# Harnessing collective intelligence

- Let people tell you what they know/what they like
- Social bookmarking



Pinterest



StumbleUpon

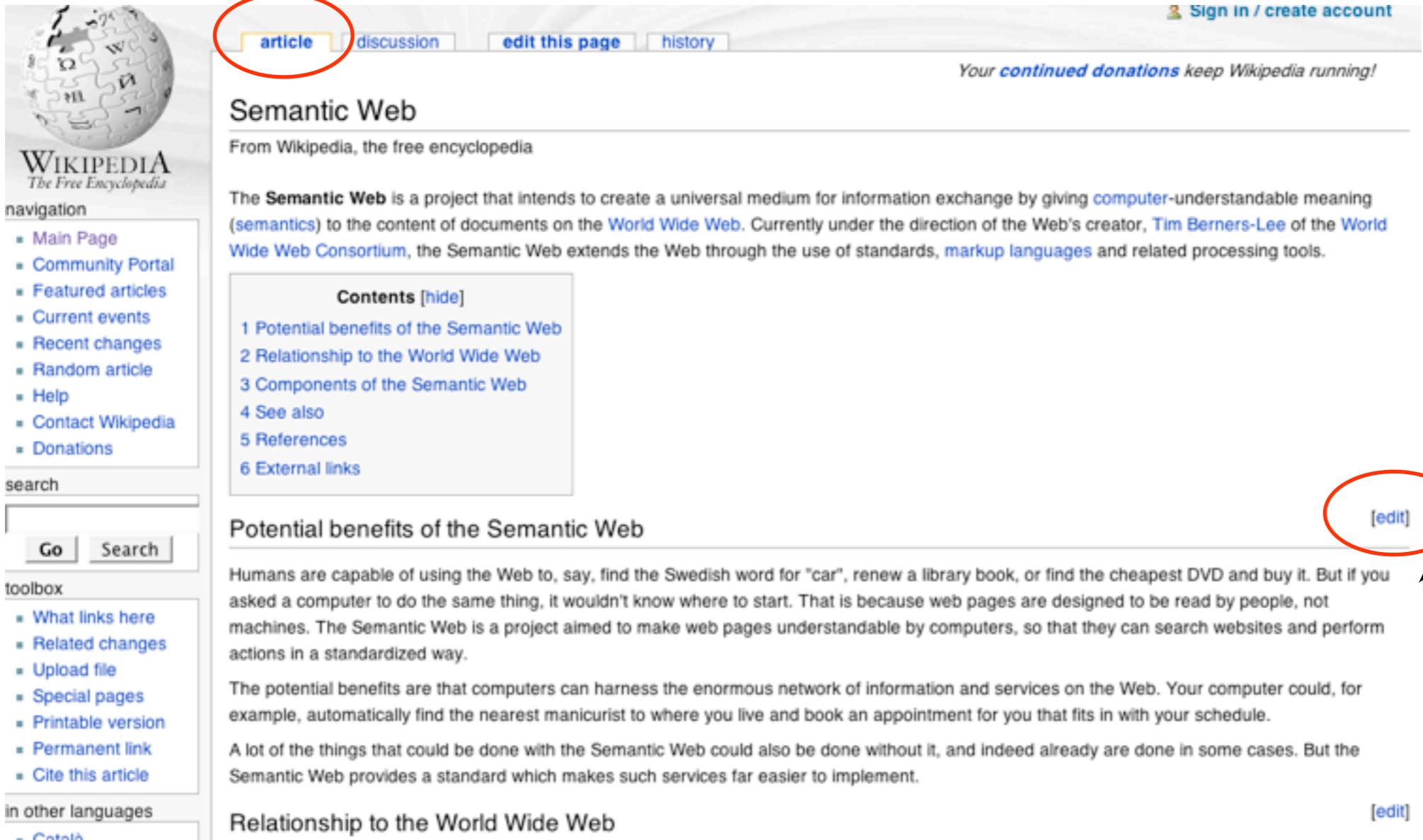
# Harnessing Collective Intelligence: Wikipedia

- Wikipedia, an online encyclopaedia based on the unlikely notion that an entry can be added by any web user, and edited by any other,
  - A radical experiment in trust
  - Eric Raymond's dictum (originally coined in the context of open source software) "with enough eyeballs, all bugs are shallow,"
- Wikipedia is already in the top 10 websites. This is a profound change in the dynamics of content creation!

**What Is Web 2.0**  
**Design Patterns and Business Models for the Next Generation of Software**

by Tim O'Reilly  
09/30/2005

# In principle: yet another encyclopaedia



Sign in / create account

article discussion edit this page history

Your continued donations keep Wikipedia running!

## Semantic Web

From Wikipedia, the free encyclopedia

The **Semantic Web** is a project that intends to create a universal medium for information exchange by giving computer-understandable meaning ([semantics](#)) to the content of documents on the [World Wide Web](#). Currently under the direction of the Web's creator, [Tim Berners-Lee](#) of the [World Wide Web Consortium](#), the Semantic Web extends the Web through the use of standards, [markup languages](#) and related processing tools.

**Contents [hide]**

- 1 Potential benefits of the Semantic Web
- 2 Relationship to the World Wide Web
- 3 Components of the Semantic Web
- 4 See also
- 5 References
- 6 External links

### Potential benefits of the Semantic Web

Humans are capable of using the Web to, say, find the Swedish word for "car", renew a library book, or find the cheapest DVD and buy it. But if you asked a computer to do the same thing, it wouldn't know where to start. That is because web pages are designed to be read by people, not machines. The Semantic Web is a project aimed to make web pages understandable by computers, so that they can search websites and perform actions in a standardized way.

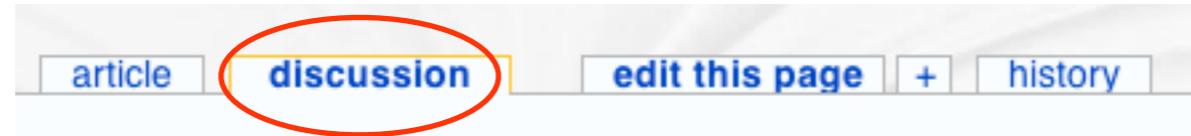
The potential benefits are that computers can harness the enormous network of information and services on the Web. Your computer could, for example, automatically find the nearest manicurist to where you live and book an appointment for you that fits in with your schedule.

A lot of the things that could be done with the Semantic Web could also be done without it, and indeed already are done in some cases. But the Semantic Web provides a standard which makes such services far easier to implement.

### Relationship to the World Wide Web

But you can edit the content!  
The collective intelligence kicks in

# Discussion



## Missing criticism

[edit]

This article seems to miss the criticisms against Semantic Web, for instance the complaint that you have to *write it twice, once for the human and once for the computer*. Secondly that Google was able to make description metatags unnecessary which suggests that the entire double-markup exercise can be solved by (meta)tagging and analysis.

Oh, and I am not putting in the NPOV tag here, that causes too much noise from troublesome editors, an example of tags having gone sour. How is that for post modern grade irony? — *The preceding unsigned comment was added by 85.164.70.218 (talk • contribs)* .

A greater concern in some critics' estimation[1] is how to find a source of revenue to pay the bandwidth bill. If 100,000 computers can continuously pull your data and understand it but cannot make purchasing decisions on their own, then there is nobody to look at the advertising.  
--[Damian Yerrick](#) (✉) 22:36, 20 February 2006 (UTC)

## First Paragraph is Incomprehensible

[edit]

By the way it was written, I imagine this wasn't written by a native speaker. I would be more than willing to help change it, but am not really sure what's trying to be said here. Could someone jump on this??

# The history to track changes (trust the users but...)



## Talk:Semantic Web

From Wikipedia, the free encyclopedia

[Revision history](#)

(Latest | Earliest) View (previous 50) (next 50) (20 | 50 | 100 | 250 | 500).

To view a previous version, click the date for that version.

Legend: (cur) = difference with current version, (last) = difference with preceding version, **m** = minor edit

[Compare selected versions](#)

- (cur) (last) ⏺ 06:50, 10 March 2006 Drito (*misc substing and minor cleanup using AWB*)
- (cur) (last) ⏺ 22:36, 20 February 2006 Damian Yerrick (*→Missing criticism - How to monetize?*)
- (cur) (last) ⏺ 22:30, 16 February 2006 Voracic (*First Paragraph is Incomprehensible*)
- (cur) (last) ⏺ 23:21, 2 February 2006 85.164.70.218 (*Missing criticism*)
- (cur) (last) ⏺ 09:04, 7 August 2005 Mjb **m** (*huh?*)

# Harnessing: The Mechanical Turk (amazon.com)

- A web service that Amazon is calling 'artificial artificial intelligence.'
- If you need a process completed that only humans can do given current technology (judgment calls, text drafting or editing, etc.),
  - you can simply make a request to the service to complete the process.
- The machine will then complete the task with volunteers, and return the results to your software."

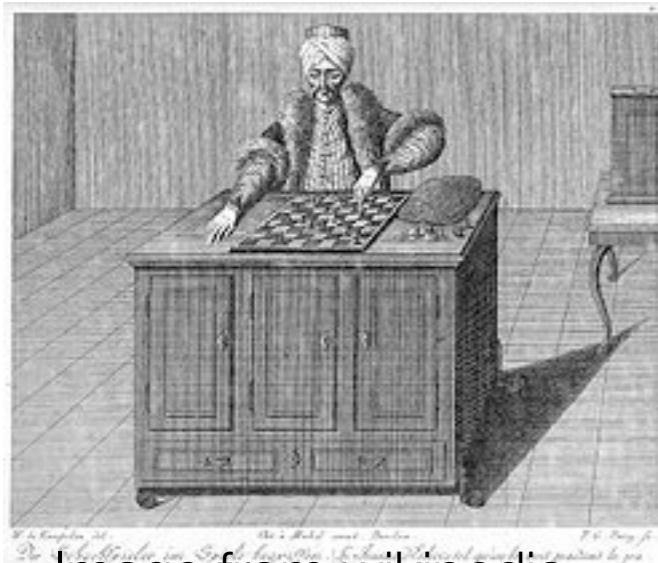


Image from wikipedia

**amazonmechanical turk**  
beta

<https://www.mturk.com/mturk/welcome>

# Other examples

- Yahoo Answers



The screenshot shows the Yahoo! Answers homepage. At the top, there's a navigation bar with the "YAHOO! ANSWERS" logo, a welcome message "Welcome, fabio.ciravegna@b...", and links for "Sign Out" and "My Account". To the right, there are links for "Answers Home", "Forum", "Blog", and "Help". Below the header, there are three main sections: "ask.", "answer.", and "discover.". The "ask." section features a large green button with a question mark icon and a placeholder text box with the instruction "Can't find it with search? Ask". A "Post Question" button is located at the bottom of this section. The "answer." section has a green button with a smiley face icon and text encouraging users to "Share knowledge", "Help others", and "Earn points". It also includes links for "What people think of Answers" and "How does it work?". The "discover." section features a green button with an exclamation mark icon and a photograph of a person wearing a headset and looking at a computer screen. It promotes becoming a better "Techie" and provides a link to "Computers & Internet". At the bottom of the page, there's a search bar with the placeholder "Search for questions:", a "Search" button, an "Advanced" link, and a "My Profile" link.

# 5 Great Ways of Harnessing Collective Intelligence

Dion Hinchcliffe

[http://www.kreeo.com/#cbok/Collective Intelligence/contents](http://www.kreeo.com/#cbok/Collective%20Intelligence/contents)  
(This blog entry has unfortunately disappeared)

- Be The Hub of A Hard To Recreate Data Source
  - This is a classic Web 2.0 concept and success here often devolves to being the first entry with an above average implementation.
    - Wikipedia, eBay, and others which are almost entirely the sum of the content their users contribute.
    - Just be careful and avoid crowded niches, like peer production news.

New features  Log

Article Discussion Read Edit View history Search



The Wikipedia content: from competitive advantage to common heritage

# 5 Great Ways of Harnessing Collective Intelligence

Dion Hinchcliffe

[http://www.kreeo.com/#cbok/Collective\\_Intelligence/contents](http://www.kreeo.com/#cbok/Collective_Intelligence/contents)

- Seek Collective Intelligence Out

- Google uses hyperlink analysis to determine the relevance of a given page and builds its own database of content which it then shares through its search engine.
  - Not only does this approach completely avoid a dependency on the ongoing kindness of strangers it also lets you build a very big content base from the outset.

Do you think  
that this ad  
will work?  
(note this is a  
spoof ad)

New features  Log in / create account

Article Discussion Read Edit View history Search 

A personal appeal from Microsoft founder Bill Gates

 **Microsoft**

**Help Bing! index pages**

From Wikipedia, the free encyclopedia

*"Contributions" redirects here. For information about the Wikipedia user contributions log, see [Help:User contributions](#).*

Fundraising or fund raising (also development) is the process of soliciting and gathering contributions as money or other resources, by requesting donations from individuals, businesses, charitable foundations, or governmental agencies (see also crowd funding). Although fundraising typically refers to efforts to gather money for non-profit organizations, it is sometimes used to refer to the identification and solicitation of investors or other sources of capital for for-profit enterprises.

## 5 ways (ctd)

- Trigger Large-Scale Network Effects
    - This is what Katrinalist and CivicSpace did and many others have done.
    - This is arguably harder to do than either of the methods above
    - Smaller examples can be found in things like the Million Dollar Pixel Page.
    - Probably not very repeatable, but when they happen, they can happen big.



**The Million Dollar Homepage - Own a piece of internet history!**

[milliondollarhomepage.com/](http://milliondollarhomepage.com/)

Block all milliondollarhomepage.com results

The **website** of Alex Tew, a 21-year-old entrepreneur, who hopes to pay his way through university by selling **1 million pixels** of internet ad space for \$1 each.

<http://haiti.ushahidi.com/>

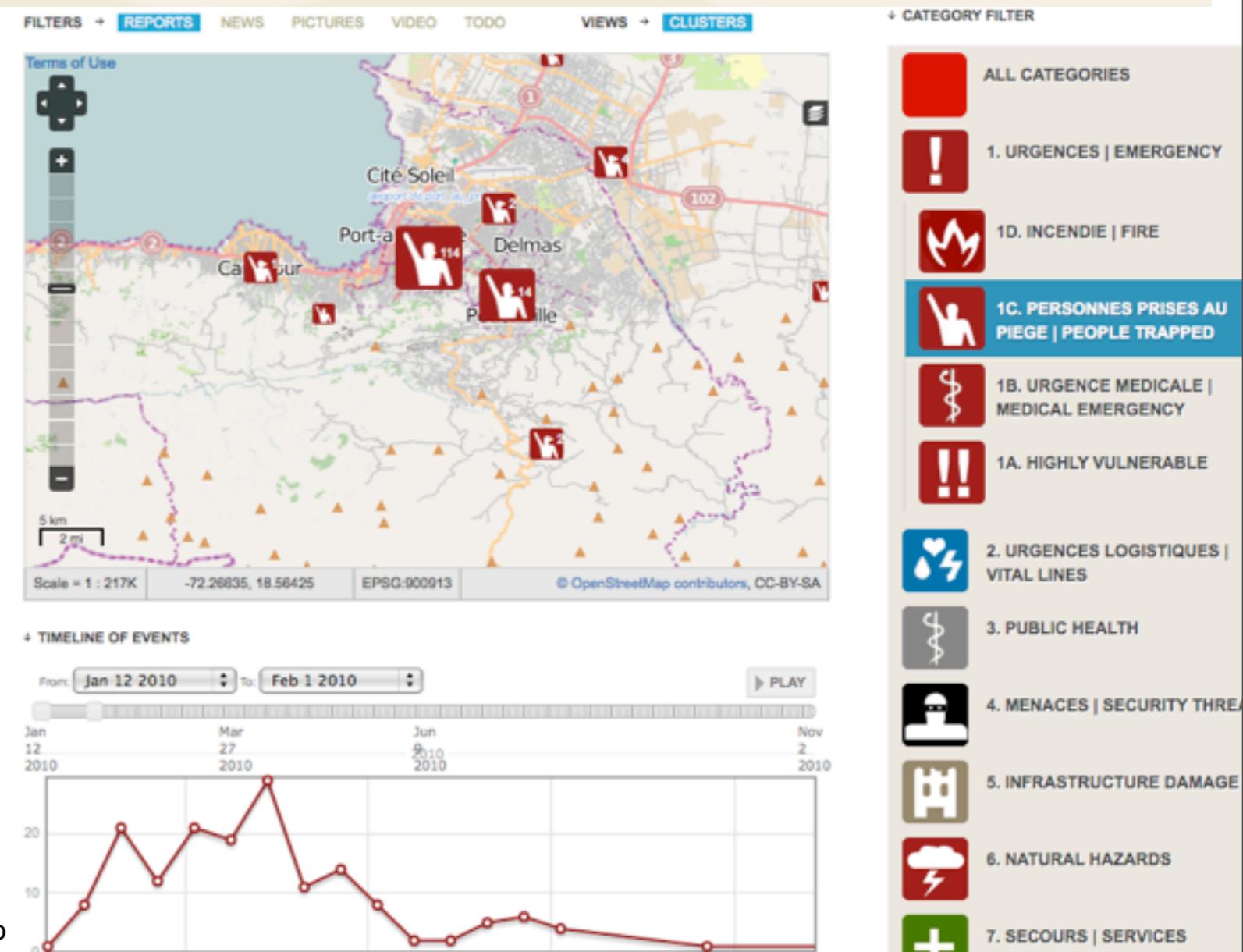


About Us Products Get Involved

Downloads >

We are a non-profit tech company that develops free and open source software for information collection, visualization and interactive mapping.

SMS, Web, Email, Radio, Phone, Twitter, Facebook, Television, List-serves, Live streams, Situation Reports



# Ushahidi Video



# Public quickly identify victims, but misidentify “shooter”

## Virginia Tech shooting shows benefits, pitfalls of social networking sites

Information about massacre spread quickly online -- but so did misinformation

By Nancy Gohring

April 18, 2007 12:00 PM ET

Comments (0)

Recommended (72)



Share

---

IDG News Service -

On Monday, many students at [Virginia Polytechnic Institute and State University](#) turned to message boards and social networking sites to try to find out what exactly was happening on campus during a shooting spree that left 33 people dead, including the perpetrator of the massacre.

But those sites sometimes spread misinformation, including an erroneous identification of the shooter that ultimately ended up being broadcast on national TV. That has raised some questions about the proper role of social networking sites and other online information sources in crises such as the one that took place at Virginia Tech.

"Social networking sites and news organizations share a couple of potential roles," said Bill Mitchell, editor of [Poynter Online](#), the Web site of the Poynter Institute, a journalism training organization in St. Petersburg, Fla. "One is to enable self-expression, and the other is to advance the story, to find out what's going on. These are roles that are sometimes in conflict."

## 5 ways (ctd)

- Provide A Folksonomy
  - It's the law of unintended uses again, something Web 2.0 design patterns strongly encourage.
  - Use real-time feedback to display tag clouds of the most popular tags and data; you'll be amazed at how much better your software works.
  - (will see later)

## 5 ways (ctd)

- Create a Reverse Intelligence Filter
  - The idea is that hyperlinks, trackbacks, and other information references can be counted and used as a reference to determine what it's important.
  - The blogosphere is the greatest example of this and sites like Memeorandum have been using this to great effect.
  - Combined with temporal filters and other techniques and you can create situation awareness engines easily.
  - It sounds similar to #2 but it's different in that it can be used with or without external data sources and is aimed not at finding but at eliding the irrelevant altogether as an active filter.

# Trusting users as co-developers

- Let people play with your data
- Let people play with your software



# Involving users as co-developers:



- Amazon sells the same products as competitors such as Barnesandnoble.com,
  - They receive the same product descriptions, cover images, and editorial content from their vendors.
- What is the difference, then?
  - Barnesandnoble.com search is likely to lead with the company's own products, or sponsored results
  - Amazon has made a science of user engagement.
    - They have an order of magnitude more user reviews, invitations to participate in varied ways on virtually every page
    - They use user activity to produce better search results.
      - Amazon always leads with "most popular", a real-time computation based on
        - Sales
        - "Flow" around products

**What Is Web 2.0**  
Design Patterns and Business Models for the Next  
Generation of Software

by Tim O'Reilly  
09/30/2005

# Involving users as developers: amazon.co.uk

Amazon.co.uk: low prices in Electronics, Books, Music, DVDs & more

Back Forward Reload Stop Home http://www.amazon.co.uk/exec/obidos/tg/stores/browse/-/welcome/468294/202-5258020-3573428 Go amazon

Getting Started Latest Headlines

**FREE Trial DVD RENTAL** Try now

**amazon.co.uk** VIEW BASKET | WISH LIST | YOUR ACCOUNT | HELP

WELCOME YOUR STORE BOOKS ELECTRONICS & PHOTO MUSIC DVD VIDEO SOFTWARE PC & VIDEO GAMES HOME & GARDEN TOYS & GAMES RENT

INTERNATIONAL HOW TO ORDER GIFT CERTIFICATES SELL YOUR STUFF HOTTEST OFFERS HARRY POTTER

SEARCH OUR SHOPS All Products GO! Make Money: by selling new or used items at [Amazon.co.uk Marketplace](#).

**STAR CHOICE**

Paperbacks -- 3 for £12 Stock up on new titles and old favourites and get 3 paperbacks for £12, including *The Apprentice* by Alan Sugar. Our Price: £3.99

**BROWSE**

- Books
- DIY & Tools
- DVD
- DVD Rental
- Electronics & Photo
- Garden & Outdoors
- Gifts
- Home & Garden
- Kitchen & Home
- Music
- PC & Video Games
- Personal Care
- Software
- Toys & Games
- Video
- Used
- Auctions

**DVD RENTAL** TRY NOW 1 Month **FREE Trial** Any 3 titles [Learn more](#)

**Must-have Toy** The award-winning [Rush Hour Traffic Jam puzzle](#) is great fun for all the family. See more in [Toys & Games](#). Our Price: £9.95

**Are you ready...for the World Cup?** Find out more

**Never Get Lost Again** No more asking for directions with the [TomTom GO 700 - Full Western Europe GPS Navigation System with Bluetooth](#).

**WHAT'S NEW**

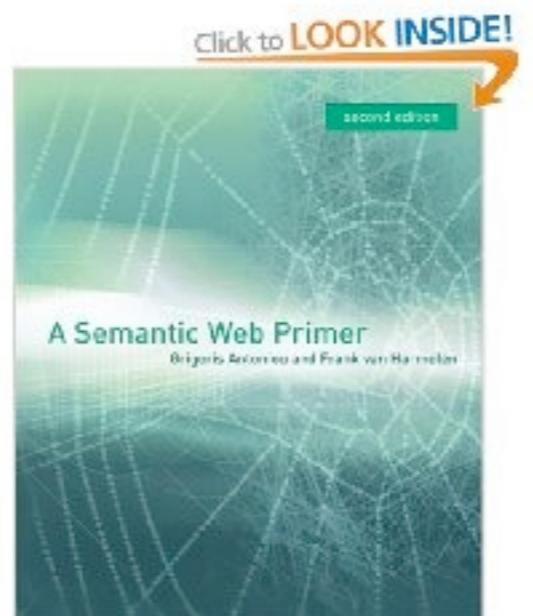
Already a customer? Sign in to see your **New for You™** recommendations.

© Fabio Ciravegna, University of Sheffield

# amazon.co.uk as Web 2.0 and As old Shop Keener

Look Inside like in a shop

Instruction to publishers  
to provide additional  
service  
(advantage to both: co-  
operation to sell more)



[See larger image](#)

[Share your own customer images](#)

Publisher: learn how customers can search inside  
this book.

Semantic Web Primer, second edition (Cooperative Information Systems Series) (Hardcover)  
by Antoniou (Author), Grigoris (Author), Harmelen (Author), Frank van (Author)  
No customer reviews yet. Be the first.

RRP: £27.95

Price: £25.15 & this item Delivered FREE in the UK with Super Saver Delivery. [See details and conditions](#)

You Save: £2.80 (10%)

In stock.

Dispatched from and sold by Amazon.co.uk. Gift-wrap available.

Only 5 left in stock--order soon (more on the way).

Want guaranteed delivery by Tuesday, March 3

23 new from £24.73 4 used from £15.00

details about the service and  
offers (standard e-commerce  
services)

more details about the product  
(standard service)

## Frequently Bought Together



Price For All Three: £88.33

[Add all three to Basket](#)

This item: Semantic Web Primer, second edition (Cooperative Information Systems Series) by Antoniou  
Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL by Dean Allemang  
Explorer's Guide to the Semantic Web by Thomas B. Passin

cluster interests based on  
past experience  
(like the old shop keeper)



## Customers Who Bought This Item Also Bought



[Semantic Web for the Working  
Ontologist: Effect... by Dean  
Allemang](#)  
 (2) £28.99



[Explorer's Guide to the Semantic  
Web by Thomas B. Passin](#)  
 (3) £34.19



[Semantic Web Technologies: Trends  
and research... by John Davies](#)  
£64.99



[Programming  
Collective  
Intelligence  
by Peter  
Norvig](#)

# Customer for customers

the judgment is shared by the users

You can judge by yourself about the quality by the comment (you may disagree)

Comments are published positive or negative (or so it seems!)

**All Customer Reviews**  
**Average Customer Review:** ★★★★  
[Write an online review and share your thoughts with other shoppers!](#)

4 of 4 people found the following review helpful:

★★★★ A high level introduction, September 5, 2005  
Reviewer: [Kaustav Bhattacharya](#) (London, UK) - [See all my reviews](#)

This book gives you a good high level introduction to the concepts and potentials of the Semantic Web. It does not go in to much detail and more it this book meant to. The author details the core concepts in a lucid manner, opening up the readers imagination to the potential application of this emerging method of organising information on the web. As well as RDF the book looks at OWL and the combination of the two emerging standards. There are many helpful references throughout the book for your further reading pleasure, many of which can be found online as well as in print.

For anyone wanting to get a good top level overview on the subject of the Semantic Web without getting bogged down with too much indepth coding detail, this book is ideal and fits very snuggly in any techie's laptop bag :-)

Was this review helpful to you?   ([Report this](#))

9 of 9 people found the following review helpful:

★★★★★ Covers all bases with clarity and in depth, September 8, 2004  
Reviewer: ["antmiller"](#) - [See all my reviews](#)

Apparently aimed at undergraduates looking at information science, this is also an invaluable introduction for technical managers transferring into this realm. The basic concepts are clearly illuminated, and all key technologies explained at a good pace: not too simple, not too rushed.

The structure of the examples is well thought out: It's generic enough to make sense in most domains, without being so general as to be meaningless.

There are many books exploring the more esoteric social and scientific edges of the semantic net, and plenty of very specialised technical tomes on each specific technology- this work bridges the gap admirably.

Was this review helpful to you?   ([Report this](#))

the community helps judge the quality of the comment

Comments on the comments

# Amazon.com: Personalised Services

 Shopping with us is safe.  
[Guaranteed.](#)

[Add to Wish List](#)  
[Add to Wedding List](#)

(We'll set one up for you)  
[View my Wish List.](#)

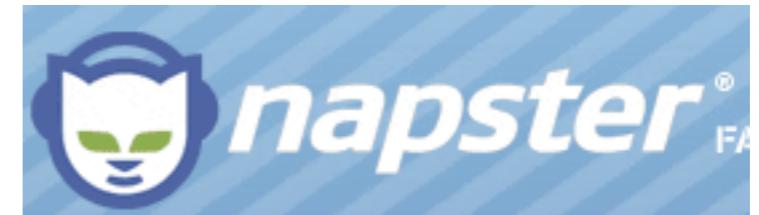
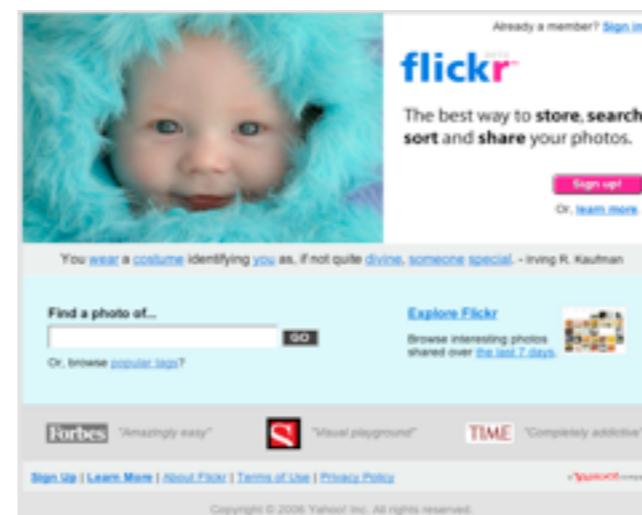
- This is my space
- I can attach my wishes to it
  - my friends can see it and send me a present
    - a mediator between people
    - like the old shop keepers who know the wishes of their customers

# Supporting community life

- Let people upload their data and share it
- Creating, supporting and free communities



Broadcast Yourself™



# The Social Web

[http://en.wikipedia.org/wiki/Social\\_Web](http://en.wikipedia.org/wiki/Social_Web)

- The social web can be described as people interlinked and interacting with engaging content in a conversational and participatory manner via the WWW
  - people are brought together through a variety of shared interests.
- Since social web applications are built to encourage communication between people, they typically emphasise some combination of the following social attributes:
  - Identity: who are you? (e.g. Facebook)
  - Reputation: what do people think you stand for? (e.g. LinkedIn)
  - Presence: where are you? (e.g. FourSquare)
  - Relationships: who are you connected with? who do you trust? (Facebook)
  - Groups: how do you organize your connections?
  - Conversations: what do you discuss with others? (e.g. Reddit)
  - Sharing: what content do you make available for others to interact with?

# The social web (ctd)

- Two types of Web sites:
  - "people focus": the person is the focus of social interaction
    - A profile is constructed by each user
    - e.g. Bebo, Facebook, and Myspace.
  - "hobby focus"
    - e.g. photography websites such as Flickr or Pinterest

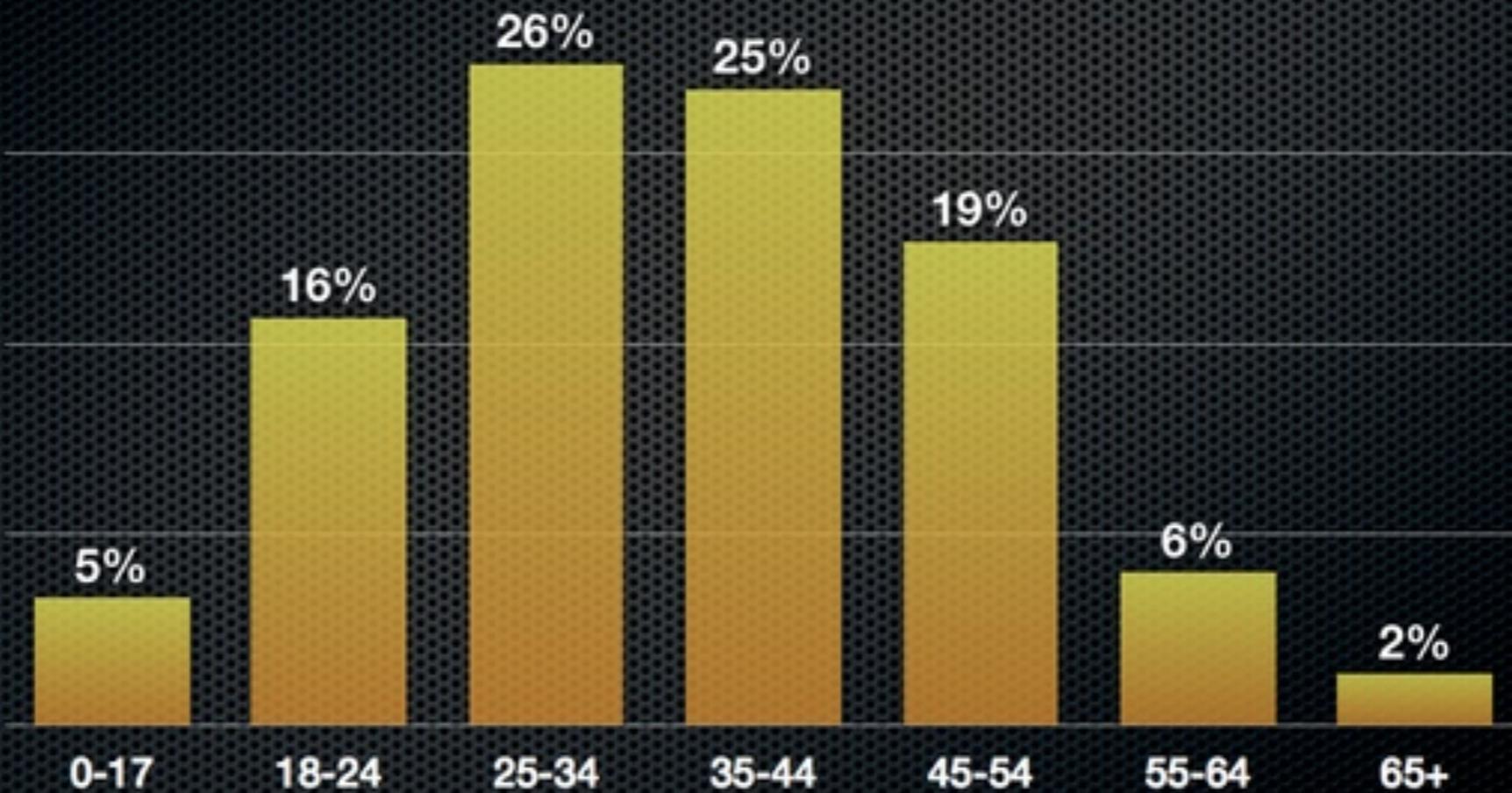
# The social web (ctd)

- The “social web” is a breed of highly interactive, participatory, conversational websites
  - blogs
  - social networking sites
  - community-edited news sites
  - services that allow sharing of bookmarks/favorites (“social bookmarking”)
  - content-sharing sites
  - discussion forums
  - newsgroups
  - public email listserves
  - wikis

# 2012

## Age distribution on social networks & online communities

Average based on the 24 sites included in this survey.

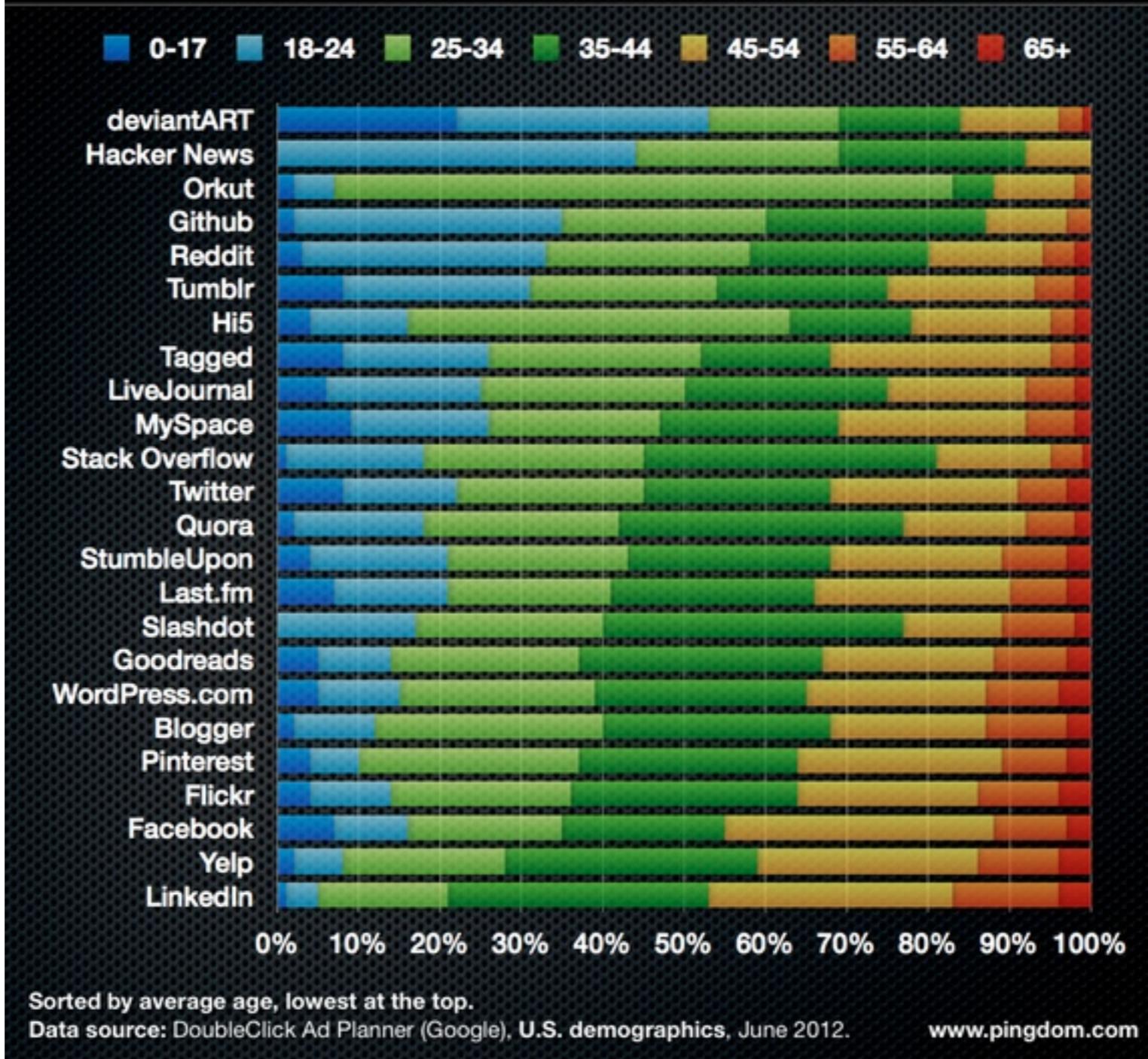


Data source: DoubleClick Ad Planner (Google), U.S. demographics, June 2012.

[www.pingdom.com](http://www.pingdom.com)

<http://royal.pingdom.com/2012/08/21/report-social-network-demographics-in-2012/>

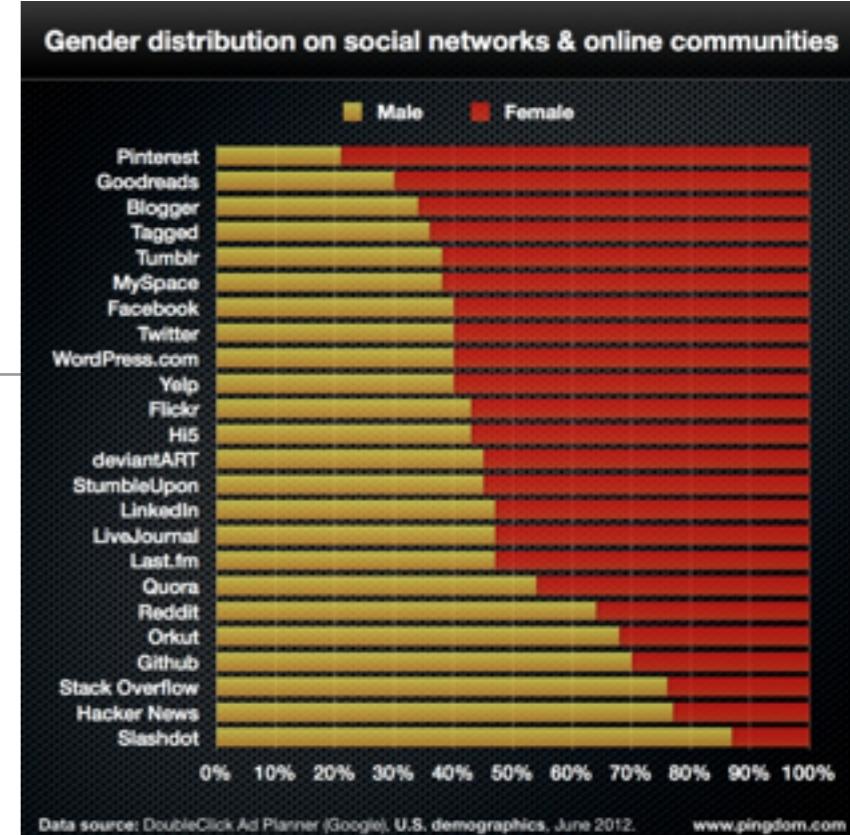
## Age distribution on social networks & online communities



- Facebook's largest community is 45-54.
- 55% of Twitter users are 35 or older.
- 63% of Pinterest users are 35 or older.
- 65% of Facebook users are 35 or older.
- 79% of LinkedIn users are 35 or older.

# Pinterest

- Pinterest is a pinboard-styled social photo sharing website.
  - The service allows users to create and manage theme-based image collections.
  - The site's mission statement is to "connect everyone in the world through the 'things' they find interesting."
- On August 16, 2011, Time magazine listed Pinterest in its "50 Best Websites of 2011" column
- August 2012: 25 million unique users,
- The fastest site in history to break through the 10 million unique visitor mark



Source: Wikipedia

# Collaborative Categorization

**What Is Web 2.0**  
Design Patterns and Business Models for the Next  
Generation of Software  
by Tim O'Reilly  
09/30/2005

- Sites like del.icio.us and Flickr have pioneered a concept that some people call "folksonomy" (in contrast to taxonomy),
  - A style of collaborative categorisation of sites using freely chosen keywords, often referred to as tags.
- Tagging allows for the kind of multiple, overlapping associations that the brain itself uses, rather than rigid categories.
- In the canonical example, a Flickr photo of a puppy might be tagged both "puppy" and "cute"--allowing for retrieval along natural axes generated user activity.

# Folksonomy

From Wikipedia

- A "folksonomy" is a collaboratively generated, open-ended labeling system that enables Internet users to categorize content such as Web pages, online photographs, and Web links.
- The freely chosen labels -- called tags -- help improve search engine's effectiveness because content is categorized using a familiar, accessible, and shared vocabulary.
- The labeling process is called tagging.
- Because folksonomies develop in Internet-mediated social environments, users can discover (generally) who created a given folksonomy tag, and see the other tags that this person created.
  - In this way, folksonomy users often discover the tag sets of another user who tends to interpret and tag content in a way that makes sense to them.
  - The result, often, is an immediate and rewarding gain in the user's capacity to find related content.

# Folksonomy

- Part of the appeal of folksonomy is its inherent subversiveness:
  - Faced with the dreadful performance of the search tools that Web sites typically provide,
    - Folksonomies can be seen as a rejection of the search engine status quo in favor of tools that are both created by the community and beneficial to the community.
- Folksonomies arise in Web-based communities where special provisions are made at the site level for creating and using tags.
- These communities are established to enable Web users to label and share user-generated content, such as photographs, or to collaboratively label existing content, such as Web sites, books, works in the scientific and scholarly literatures, and blog entries.

# Blogging and the Wisdom of Crowds

- One of the most highly touted features of the Web 2.0 era is the rise of blogging.
- Personal home pages have been around since the early days of the web, and the personal diary and daily opinion column around much longer than that, so just what is the fuss all about?
- At its most basic, a blog is just a personal home page in diary format. But as Rich Skrenta notes, the chronological organization of a blog "seems like a trivial difference, but it drives an entirely different delivery, advertising and value chain."

**What Is Web 2.0**  
**Design Patterns and Business Models for the Next Generation of Software**

by Tim O'Reilly  
09/30/2005

© Fabio Ciravegna, University of Sheffield

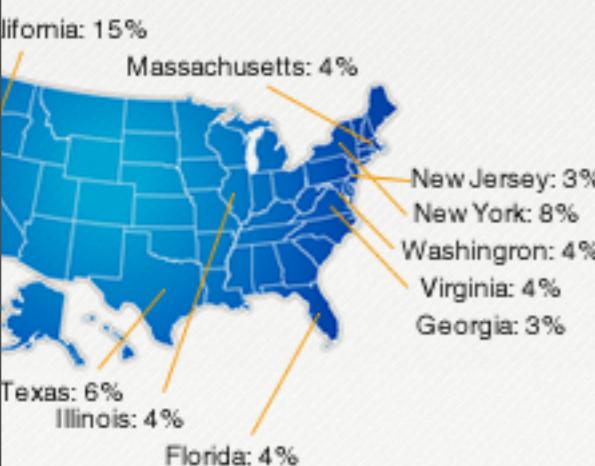
The

# The Blogosphere (2011)

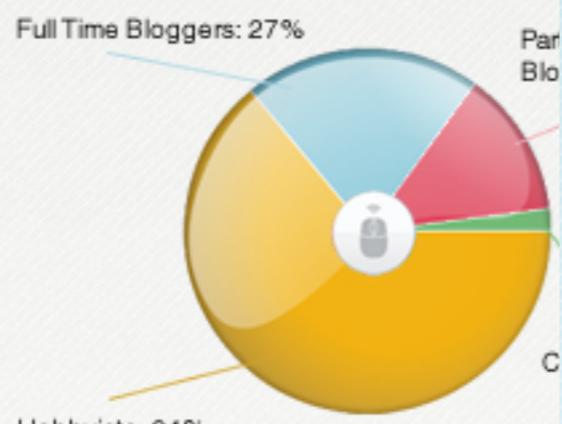
## Number of Blogs according to Technorati & Blogr



## Top states for Bloggers in US



## Bloggers Type



Demographic



2/3 male 1/3 Females

- 18-24 years 7%
- 25-34 years 30%
- 35-44 years 27%
- 45-49 years 11%
- 50-54 years 9%
- 55-64 years 12%
- 65+ years 4%



USA - 49%



Canada & Mexico - 7%



EU - 29%



South America - 3%



Asia Pacific - 12%

## How long you've been blogging



More than 6 years



4-6 years



2-4 years



1-2 years



less than 1 year

## Common Methods to Measure Success

- 😊 66% Personal Satisfaction
- 🌐 34% Social media sites
- \$ 11% Revenue Generated
- 🔗 39% Backlinks

- 👤 50% Unique Visitors
- 💬 51% Number of posts and comments
- 💻 14% Coverage on Traditional Media
- 📠 10% Business Leads

11% of Bloggers earn their primary income from Blogging

## Upcoming frequency per month

- One per month - 8%
- 2-4 times - 15%
- 4-12 times - 55%
- 30-60 times - 20%
- 90-120 times - 4%
- 150-300 times - 3%

## Time Spent on Blogging each week

- Less than 1 hour - 32%
- 1-3 hours - 28%
- 3-5 hours - 16%
- 5-10 hours - 12%
- 10-20 hours - 7%
- 20-40 hours - 4%

## Why bloggers Blog?

- 35% to write about areas of interest
- 26% to share expertise
- 10% to attract new clients
- 7% to connect with like-minded people
- 4% to keep in touch with friends and family
- 2% to get featured on traditional media sites

<http://www.invesp.com/blog/business/how-big-is-blogosphere.html>

# RSS Feeds

- One of the things that has made a difference is a technology called RSS.
- RSS is the most significant advance in the fundamental architecture of the web since early hackers realized that CGI could be used to create database-backed websites.
- RSS allows someone to link not just to a page, but to subscribe to it, with notification every time that page changes.
- Some call it the "live web".
  - Dynamic websites" (i.e., database-backed sites with dynamically generated content) replaced static web pages well over ten years ago
  - What's dynamic about the live web are not just the pages, but the links.
  - A link to a weblog is expected to point to a perennially changing page, with "permalinks" for any individual entry, and notification for each change.
  - An RSS feed is thus a much stronger link than, say a bookmark or a link to a single page.

# RSS Feeds (2)

- RSS also means that the web browser is not the only means of viewing a web page.
  - While some RSS aggregators, such as Bloglines, are web-based, others are desktop clients, and still others allow users of portable devices to subscribe to constantly updated content.
- RSS is now being used to push not just notices of new blog entries, but also all kinds of data updates, including stock quotes, weather data, and photo availability.

# Twitter (2006)



- Twitter is a social networking and micro-blogging service that allows its users to send and read other users' updates (known as tweets),
  - text-based posts of up to 140 characters in length.
  - Updates are displayed on the user's profile page and delivered to other users who have signed up to receive them.

<http://en.wikipedia.org/wiki/Twitter>

# Product Trends

**What Is Web 2.0**  
Design Patterns and Business Models for the Next  
Generation of Software

by Tim O'Reilly  
09/30/2005

- Services, not packaged software, with cost-effective scalability
- Control over unique, hard-to-recreate data sources that get richer as more people use them
- Software above the level of a single device
- Lightweight user interfaces, development models, AND business models

# From Proprietary Software to Proprietary Data

- Every significant internet application to date has been backed by a specialized database:
  - Google's web crawl and indexing
  - Yahoo!'s directory (and web crawl)
  - Amazon's database of products and user contributions
  - eBay's database of products and sellers,
- Database management is a core competency of Web 2.0 companies, so much so that we have sometimes referred to these applications as "infoware" rather than merely software.

**What Is Web 2.0**  
Design Patterns and Business Models for the Next Generation of Software

by Tim O'Reilly  
09/30/2005

# Data Ownership = Success?

**What Is Web 2.0**  
Design Patterns and Business Models for the Next  
Generation of Software  
by Tim O'Reilly  
09/30/2005

- A number of cases where control over the database has led to market control and outsized financial returns
  - Look at the copyright notices at the base of every map served by MapQuest, maps.yahoo.com, maps.msn.com, or maps.google.com,
    - you'll see the line "Maps copyright NavTeq, TeleAtlas,"
    - for the new satellite imagery services, "Images copyright Digital Globe."
- These companies made substantial investments in their databases
  - NavTeq alone reportedly invested \$750 million to build their database of street addresses and directions.

# NavTeq

- NavTeq has gone so far as to imitate Intel's familiar Intel Inside logo:
  - Cars with navigation systems bear the imprint, "NavTeq Onboard."
  - Data is indeed the Intel Inside of these applications, a sole source component in systems whose software infrastructure is largely open source or otherwise commodified.

# The failure of MapQuest

**What Is Web 2.0**  
Design Patterns and Business Models for the Next  
Generation of Software

by Tim O'Reilly  
09/30/2005

- A failure to understand the importance of owning an application's core data will eventually undercut its competitive position.
  - MapQuest pioneered the web mapping category in 1995
  - When Yahoo!, and then Microsoft, and most recently Google, decided to enter the market, they were easily able to offer a competing application simply by licensing the same data

# How to make use of Data: Amazon.com

- Instead Amazon.com
  - Original database came from ISBN registry provider R.R. Bowker.
    - Like competitors such as Barnesandnoble.com,
  - But Amazon
    - Enhanced the data,
      - adding publisher-supplied data such as cover images, table of contents, index, and sample material.
    - Harnessed their users to annotate the data
- Now Amazon (not Bowker) is the primary source for bibliographic data on books, a reference source for scholars and librarians as well as consumers.
- Imagine if MapQuest had done the same thing, harnessing their users to annotate maps and directions, adding layers of value. It would have been much more difficult for competitors to enter the market just by licensing the base data.

**What Is Web 2.0**  
**Design Patterns and Business Models for the Next Generation of Software**

by [Tim O'Reilly](#)  
09/30/2005

# Google Maps

- Google Maps provides a living laboratory for the competition between application vendors and their data suppliers.
- Google's lightweight programming model has led to the creation of numerous value-added services in the form of mashups that link Google Maps with other internet-accessible data sources.
- Google has taken the role of data source away from Navteq and inserted themselves as a favored intermediary.

# Other Issues

- The Long Tail

Small sites make up the bulk of the internet's content; narrow niches make up the bulk of internet's the possible applications. Therefore: Leverage customer-self service and algorithmic data management to reach out to the entire web, to the edges and not just the center, to the long tail and not just the head.
- Network Effects by Default

Only a small percentage of users will go to the trouble of adding value to your application. Therefore: Set inclusive defaults for aggregating user data as a side-effect of their use of the application.
- Some Rights Reserved. Intellectual property protection limits re-use and prevents experimentation.
  - Therefore: When benefits come from collective adoption, not private restriction, make sure that barriers to adoption are low.
  - Follow existing standards, and use licenses with as few restrictions as possible. Design for "hackability" and "remixability."

# The Long Tail

- Overture and Google's success came from an understanding of what Chris Anderson refers to as "the long tail," the collective power of the small sites that make up the bulk of the web's content.
  - DoubleClick's offerings require a formal sales contract, limiting their market to the few thousand largest websites.
  - Overture and Google figured out how to enable ad placement on virtually any web page.
    - What's more, they eschewed publisher/ad-agency friendly advertising formats such as banner ads and popups in favor of minimally intrusive, context-sensitive, consumer-friendly text advertising.
- The Web 2.0 lesson: leverage customer-self service and algorithmic data management to reach out to the entire web, to the edges and not just the center, to the long tail and not just the head.

# A conclusion from 2005!

- It is for this reason that a recent ZDnet editorial [concluded that Microsoft won't be able to beat Google](#): "Microsoft's business model depends on everyone upgrading their computing environment every two to three years. Google's depends on everyone exploring what's new in their computing environment every day."
- See Google Apps (spreadsheets, etc.)

**What Is Web 2.0**  
**Design Patterns and Business Models for the Next Generation of Software**  
by Tim O'Reilly  
09/30/2005

Welcome to Google Apps

Change Language: English (US) ▾

**Google innovation. Powerful solutions. Low cost.**

**New!** Try Google Apps Premier Edition for free through April 30th, 2007.

With Google Apps, you can give your employees the next-generation communication and collaboration tools they need to manage electronic communications, share and publish information and stay connected while on the go.

Whether you're looking to [transition from or complement](#) your existing messaging infrastructure with more advanced collaboration tools, Google Apps gives you full control while requiring minimal investment. Google Apps can also make it easy to meet [deskless employees'](#) email and calendaring needs.

Best of all, it's all hosted by Google, so there's no hardware or software to download, install or maintain. With Google, you can afford to provide each and every employee with the tools they need to succeed.

**Compare Editions and Sign Up**



**Communicate and connect**

 **Gmail**  
Email with 10 GB of storage per custom email account, mail search tools and integrated chat.

 **Google Talk**  
Free text and voice calling around the world.

 **Google Calendar**  
Coordinate meetings and company events with sharable calendars.

Learn more about the [services and tools](#) available though Google Apps.

**Collaborate and publish**

 **Start Page**  
Access your inbox, calendar, docs and company info, plus search the web from one place.

 **Docs & Spreadsheets**  
Create, share and collaborate on documents in real-time.

 **Page Creator**  
Easily create and publish web pages.

**Manage your services**

 **Control Panel**  
Manage your domain and user accounts online.

 **Extensibility APIs**  
Integrate with your existing IT systems or 3rd party solutions.

 **Help and support**  
24/7 assistance, including phone support.

# Some techy details



Google maps  
UK

AJAX.org

# Towards a different Client-Server Architecture

- Classic web application model
  - Most user actions in the interface trigger an HTTP request back to a web server.
  - The server does some processing — retrieving data, crunching numbers, talking to various legacy systems — and then returns an HTML page to the client.
- This approach makes a lot of technical sense, but it doesn't make for a great user experience.
  - While the server is doing its thing, what's the user doing?
  - That's right, waiting.
  - And at every step in a task, the user waits some more.
- Obviously, if we were designing the Web from scratch for applications, we wouldn't make users wait around. Once an interface is loaded, why should the user interaction come to a halt every time the application needs something from the server? In fact, why should the user see the application go to the server at all?

Ajax: A New Approach to Web Applications

by [Jesse James Garrett](#)

<http://adaptivepath.com/publications/essays/archives/000385.php>

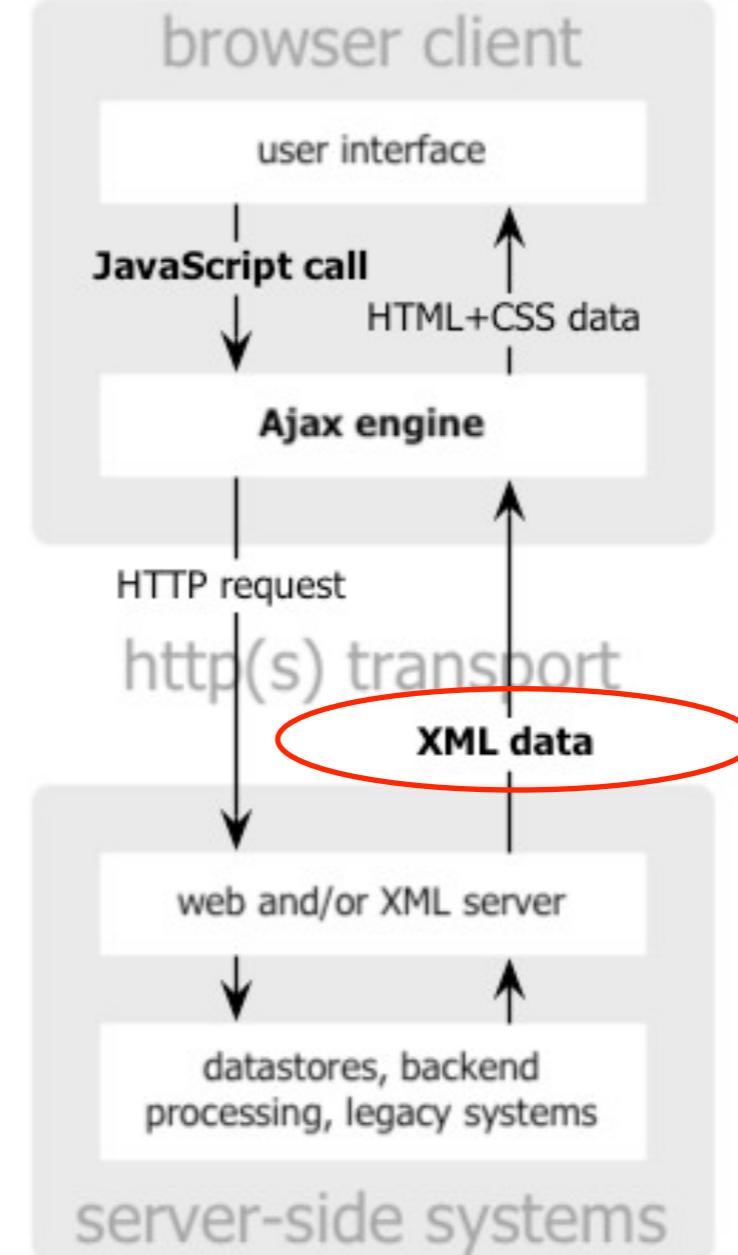
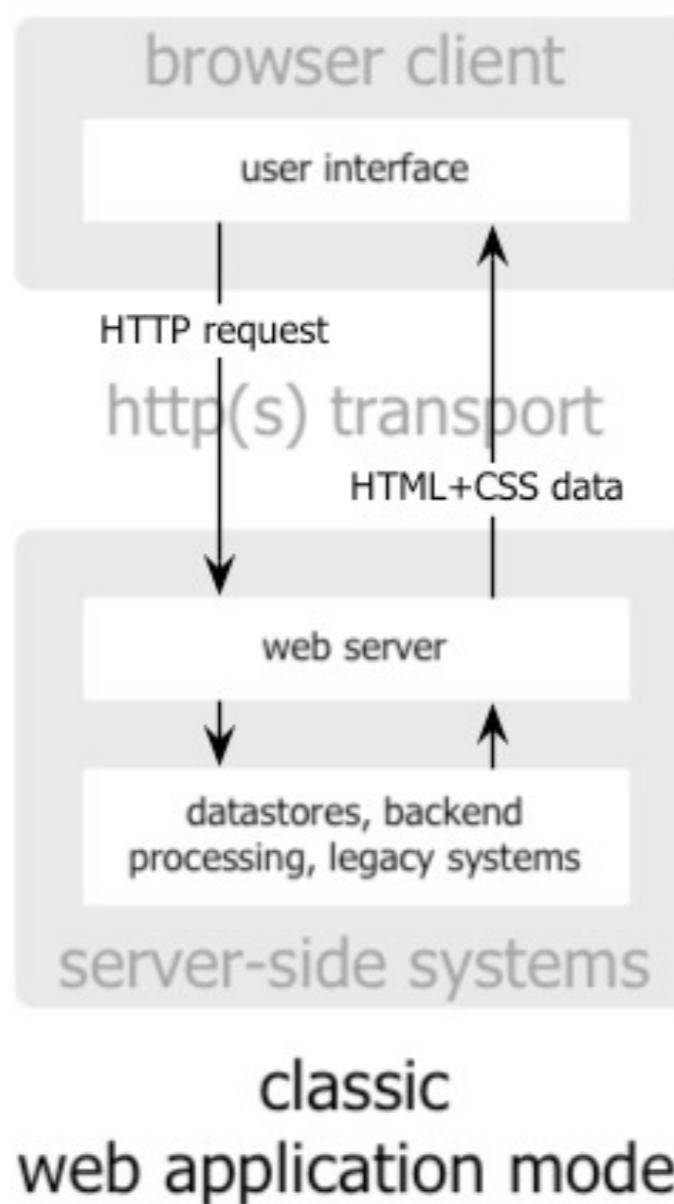
# Ajax is different

Ajax: A New Approach to Web Applications

by [Jesse James Garrett](#)

<http://adaptivepath.com/publications/essays/archives/000385.php>

- An Ajax application eliminates the start-stop-start-stop nature of interaction on the Web by introducing an intermediary — an Ajax engine — between the user and the server
- At the start of the session, the browser loads an Ajax engine (not just a Web page)
  - Written in JavaScript and usually tucked away in a hidden frame.
  - This engine is responsible for both rendering the interface the user sees and communicating with the server on the user's behalf.
- The Ajax engine allows the user's interaction with the application to happen asynchronously — independent of communication with the server. So the user is never staring at a blank browser window and an hourglass icon, waiting around for the server to do something.



# What does the Ajax engine do?

- Every user action that normally would generate an HTTP request takes the form of a JavaScript call to the Ajax engine instead.
- Any response to a user action that doesn't require a trip back to the server — such as simple data validation, editing data in memory, and even some navigation — the engine handles on its own.
- If the engine needs something from the server in order to respond — if it's submitting data for processing, loading additional interface code, or retrieving new data — the engine makes those requests asynchronously, usually using XML, without stalling a user's interaction with the application.
- Example:
  - visualise a Google Map <http://maps.google.com/>
  - grab the top right corner of the map
  - pull towards left bottom corner

# Success of Ajax

- Google is making a huge investment in developing the Ajax approach.
- All of the major products Google has introduced over the last years — [Gmail](#), the latest beta version of [Google Groups](#), [Google Suggest](#), and [Google Maps](#) — are Ajax applications
- Others are following suit: many of the features that people love in [Flickr](#) depend on Ajax.

*We're entering an unprecedented period of user interface innovation, as web developers are finally able to build web applications as rich as local PC-based applications.*

**What Is Web 2.0**  
**Design Patterns and Business Models for the Next Generation of Software**

by [Tim O'Reilly](#)  
09/30/2005

# Ajax FAQ

- Do Ajax applications always deliver a better experience than traditional web applications?
- A. Not necessarily. Ajax gives interaction designers more flexibility.
  - However, the more power we have, the more caution we must use in exercising it.
  - We must be careful to use Ajax to enhance the user experience of our applications, not degrade it.

# A little Ajax program

- Ajax uses a programming model with display and events. These events are user actions, they call functions associated to elements of the web page.
- Interactivity is achieved with forms and buttons.
- To get data on the server, XMLHttpRequest provides two methods:
  - open: create a connection.
  - send: send a request to the server.
- Data returned by the server will be found in the attributes of the XMLHttpRequest object:
  - responseXml for an XML file or
  - responseText for a plain text.

# The XMLHttpRequest object

- Allows to interact with the servers, thanks to its methods and attributes

## Attributes

readyState	the code successively changes value from 0 to 4 that means for "ready".
status	200 is OK 404 if the page is not found.
responseText	holds loaded data as a string of characters.
responseXml	holds an XML loaded file, DOM's method allows to extract data.
onreadystatechange	property that takes a function as value that is invoked when the readystatechange event occurs.

## Methods

open(mode, url, boolean)	mode: type of request, GET or POST url: the location of the file, with a path. boolean: true (asynchronous) / false (synchronous). optionally, a login and a password may be added to arguments.
send("string")	null for a GET command.

# First step: create an instance

```
if (window.XMLHttpRequest) // Object of the current windows
{
  xhr = new XMLHttpRequest(); // Firefox, Safari, ...
}
else
if (window.ActiveXObject) // ActiveX version
{
  xhr = new ActiveXObject("Microsoft.XMLHTTP"); // Internet Explorer
}
```

## Second Step: wait for the response

```
xhr.onreadystatechange = function() { // instructions to process the response };

if (xhr.readyState == 4)
{
    // Received, OK
} else
{
    // Wait...
}
```

## Third step: make the request

- Two methods of XMLHttpRequest are used:
  - open: command GET or POST, URL of the document, true for asynchronous.
  - send: with POST only, the data to send to the server.
- The request below read a document on the server.
  - `xhr.open('GET', 'http://www.xul.fr/somefile.xml', true);`
  - `xhr.send(null);`

# Complete Code

```

<html><head>
<script>
function submitForm() {
    var xhr;
    try { xhr = new ActiveXObject('Msxml2.XMLHTTP'); }
    catch (e) {
        try { xhr = new ActiveXObject('Microsoft.XMLHTTP'); }
        catch (e2) {
            try { xhr = new XMLHttpRequest(); }
            catch (e3) { xhr = false; }
        }
    }
    xhr.onreadystatechange = processChange; ←
    xhr.open(POST, "data.txt", true);
    xhr.setRequestHeader("Content-type","application/x-www-form-urlencoded");
    var namevalue=encodeURIComponent(document.getElementById("name").value)
    var agevalue=encodeURIComponent(document.getElementById("age").value)
    var parameters="name="+namevalue+"&age="+agevalue
    xhr.send(parameters)
}

function processChange() {
    if(xhr.readyState == 4) {
        if(xhr.status == 200)
            document.ajax.responseText;
        else
            document.ajax.responseText = "Error code " + xhr.status;
    }
}

```

Create object

What to do when server responds

(processChange is a function that will be called)

Post to server

When the function terminates, the request has been sent and the browser is free to do something else

When server responds, do something

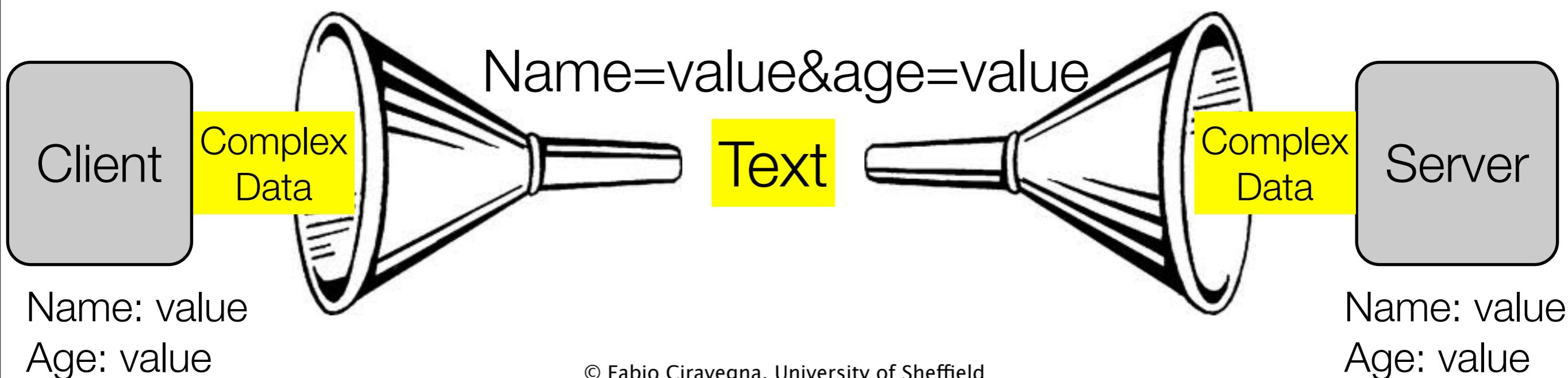
# Client Server Communication

JSON

# Client-Server communication

- Different environments
  - E.g. JavaScript on browser and Java on server
  - Difficult communication
- Solution?
  - Serialisation/de-serialisation of data into text

This approach is really cumbersome



# Example: National Rail Enquiry

<http://www.nationalrail.co.uk/>



**National Rail Enquiries**

Register now for instant access to your favourite journeys  
Already registered? [Sign in now](#).

**Let's go!**

Home Train times & tickets Stations & on train Changes to train times Hotels Search site **Search** **YAHOO!**

**i** Travel news > Prepare for winter weather with our information feeds

### Find my train times & fares

From  to  Leaving  Today  at   :

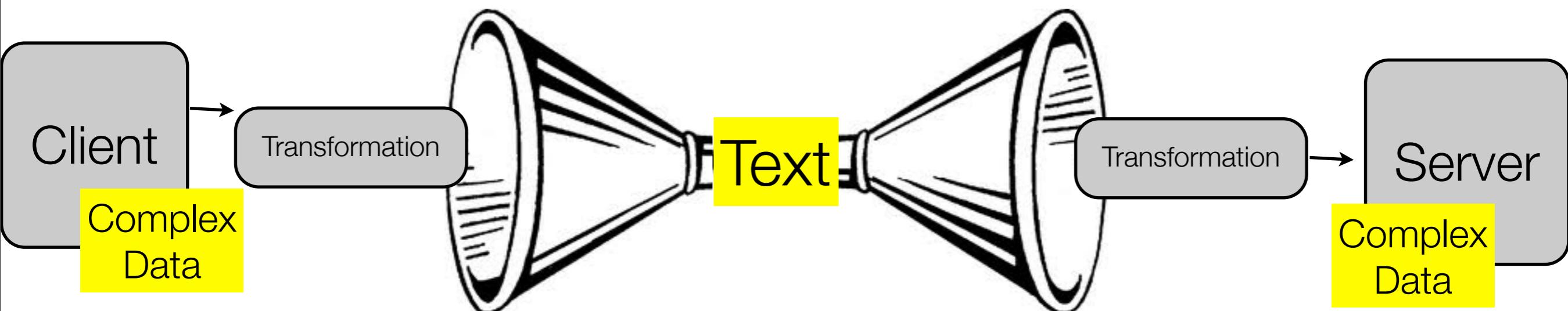
Remove return journey Leaving    at   :   **GO**

Advanced search   Show only fastest trains

# Traditional solution

- Client and server exchange
  - Client to Server (POST)
    - fromStation="Shf"&toStation="MncAir"&dateOut="Today"&hourOut="9"&minOur ="15"&dateIn="21/03/2012"&....
  - Server to Client: HTML string

This approach is really cumbersome



Not for turning the data into text  
but because of the way we encode it

# JSON

<http://www.json.org/>

- JSON (JavaScript Object Notation) is a lightweight data-interchange format.
  - It is easy for humans to read and write. It is easy for machines to parse and generate.
  - It is based on a subset of the JavaScript Programming Language
  - JSON is a text format that is completely language independent but uses conventions that are familiar to programmers of the C-family of languages, including C, C++, C#, Java, JavaScript, Perl, Python, and many others.
  - These properties make JSON an ideal data-interchange language.

# JSON (2)

- JSON is built on two structures:
  - A collection of name/value pairs. In various languages, this is realized as an object, record, struct, dictionary, hash table, keyed list, or associative array.
  - An ordered list of values. In most languages, this is realized as an array, vector, list, or sequence.
  - These are universal data structures. Virtually all modern programming languages support them in one form or another. It makes sense that a data format that is interchangeable with programming languages also be based on these structures.

# JSON

- Gson is a Java library that can be used to convert Java Objects into their JSON representation. It can also be used to convert a JSON string to an equivalent Java object. Gson can work with arbitrary Java objects including pre-existing objects that you do not have source-code of
  - Download the gson library in order to use it (it is not in the standard java distribution)

<http://code.google.com/p/google-gson/>

# Serialisation (toGson)

- Serialisation:
- in the context of data storage and transmission, serialization is the process of converting a data structure or object state into a format that can be stored (for example (...) transmitted across a network connection link) and "resurrected" later in the same or another computer environment (wikipedia)

```
/* create Gson object */
Gson gson = new Gson();
/* create the object to serialise (any Java object)*/
class BagOfPrimitives {
    private int value1 = 1;
    private String value2 = "abc";
    private transient int value3 = 3;
    BagOfPrimitives() {
        // no-args constructor
    }
}
BagOfPrimitives obj = new BagOfPrimitives();
String json = gson.toJson(obj);
```

# Deserialisation

- The opposite operation: extracting a data structure from a series of bytes

```
BagOfPrimitives obj2 =
```

```
    gson.fromJson(json, BagOfPrimitives.class);
```

Json object

(it does not need to be Gson)

Typically the first parameter of a doGet/doPost

Expected Object class

```
public class TestServlet extends HttpServlet {  
    public void doGet(HttpServletRequest req, HttpServletResponse res)  
        throws IOException, ServletException {
```

# JavaScript side

<http://www.json.org/js.html>

- Deserialisation

```
var myObject = JSON.parse(myJSONtext, reviver);
```

- The optional *reviver* function will be called for every key and value at every level of the final result.
  - Each value will be replaced by the result of the reviver function.
  - This can be used e.g. to transform date strings into Date objects.

```
myData = JSON.parse(text, function (key, value) {  
    var type;  
    if (value && typeof value === 'object') {  
        type = value.type;  
        if (typeof type === 'string' && typeof window[type]  
            'function') {  
            return new (window[type])(value);  
        }  
    }  
    return value;
```

# JavaScript side (2)

- Serialisation

```
var myJSONText = JSON.stringify(myObject, replacer);
```

- The stringifier method can take an optional replacer function.
  - called after the toJSON method on each of the values in the structure.
  - It will be passed each key and value as parameters, and this will be bound to object holding the key.
  - The value returned will be stringified.

```
function replacer(key, value) {  
    if (typeof value === 'number' && !isFinite(value)) {  
        return String(value);  
    }  
    return value;  
}
```

# Why use it?

- Client server architecture can now return data structures as opposed to HTML code
  - Client can send complex objects (as opposed to just variable-value pairs)
  - Client is no longer passive: now it interprets the code and displays it as required





National Rail Enquiries

Register now for instant access to your favourite journeys  
Already registered? [Sign in now](#).

Home Train times & tickets Stations & on train Changes to train times Hotels

Search site Search YAHOO!

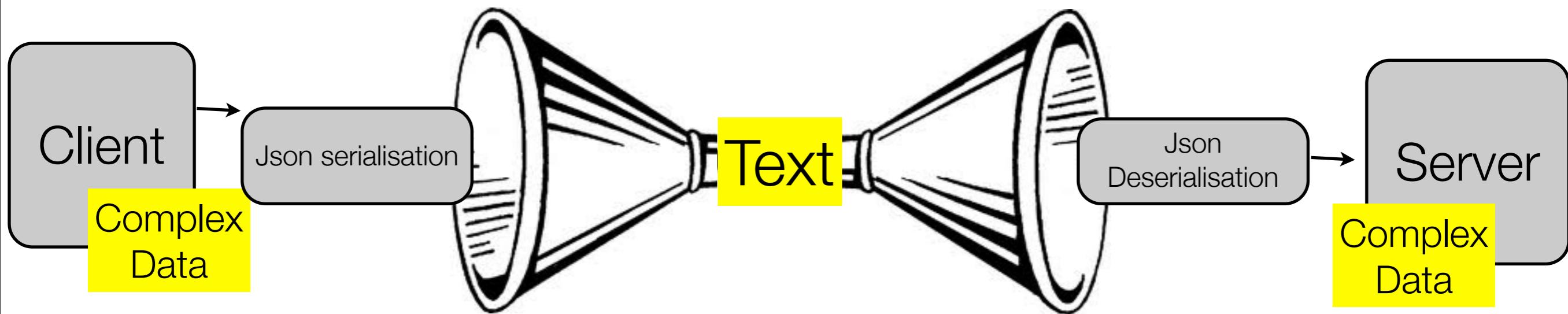
Travel news > Prepare for winter weather with our information feeds

Find my train times & fares

From Sheffield to Manchester Airport Leaving Today at 09:15

Remove return journey Leaving 21/03/2012 at 11:15 GO

Advanced search Passengers: 1 Adult Show only fastest trains ?



We do not need to find a clever way to encode the data structure. Json does it for you

# Twitter API



# Twitter API

<http://net.tutsplus.com/tutorials/other/diving-into-the-twitter-api/>

- There are three separate Twitter APIs actually.
  - The normal REST based API
    - methods constitute the core of the Twitter API, and are written by Twitter itself. It allows other developers to access and manipulate all of Twitter's main data.
    - You'd use this API to do all the usual stuff you'd want to do with Twitter including retrieving statuses, updating statuses, showing a user's timeline, sending direct messages and so on.
  - The Search API
    - Lets you look beyond you and your followers. You need this API if you are looking to view trending topics and so on.
  - The Stream API
    - lets developers sample huge amounts of real time data. Since this API is only available to approved users, we aren't going to go over this today.
- Public data can be freely accessed without an API key. When requesting private data and/or user specific data, Twitter requires authentication

## The API (ctd)

- There are limits to how many calls and changes you can make in a day <http://dev.twitter.com/pages/rate-limiting>
  - API usage is rate limited with additional fair use limits to protect Twitter from abuse.
- The API is entirely HTTP-based
  - Methods to retrieve data from the Twitter API require a GET request. Methods that submit, change, or destroy data require a POST.
  - API Methods that require a particular HTTP method will return an error if you do not make your request with the correct one.
  - HTTP Response Codes are meaningful
- The API presently supports the following data formats: XML, JSON, and the RSS and Atom syndication formats, with some methods only accepting a subset of these formats.

# REST API Methods

<https://dev.twitter.com/docs/api/1.1>

- Timeline Methods
  - statuses/public\_timeline
  - statuses/home\_timeline
  - statuses/friends\_timeline
  - statuses/user\_timeline
  - statuses/mentions
  - statusesretweeted\_by\_me
  - statusesretweeted\_to\_me
  - statusesretweets\_of\_me
- And several others!!!!

# Interacting with Twitter in Java

<http://twitter4j.org>

- Twitter4J is an unofficial Java library for the Twitter API.
  - You can easily integrate Java application with the Twitter service
  - Twitter4J is featuring:
    - 100% Pure Java - works on any Java Platform version 1.4.2 or later
    - Android platform and Google APP Engine ready
    - Zero dependency : No additional jars required
    - Built-in OAuth support
    - Out-of-the-box gzip support
- Just download and add its jar file to the application classpath.

# An Example

- Get tweets from people in Sheffield about Sheffield
  - People in Sheffield == geolocated in Sheffield
  - About Sheffield == using #Sheffield
- A number of examples at  
<https://github.com/yusuke/twitter4j/tree/master/twitter4j-examples/src/main/java/twitter4j/examples>

```
package tryno.server;
import twitter4j.*;
import java.util.List;
public class GetTweetByLocation {
```

```
public String getSimpleTimeLine(){
    Twitter twitter = new TwitterFactory().getInstance();
    String resultString= "";
    try {
        // it creates a query and sets the geocode
        //requirement
        Query query= new Query("#sheffield");
        query.setGeoCode(new GeoLocation(53.383, -1.483), 2,
                         Query.KILOMETERS);
        //it fires the query
        QueryResult result = twitter.search(query);
```

(ctd)

(ctd from previous slide)

```
//it cycles on the tweets
List<Tweet> tweets = result.getTweets();
for (Tweet tweet : tweets) {
    ///gets the user
    User user = twitter.showUser(tweet.getFromUser());
    Status status= (user.isGeoEnabled())?user.getStatus():null;
    if (status==null)
        resultString+="@ " + tweet.getFromUser() + " (" +
                    + user.getLocation()
                    + ") - " + tweet.getText();
    else resultString+="@ " + tweet.getFromUser() + " (" + ((status!
=null&&status.getGeoLocation()!=null)?
status.getGeoLocation().getLatitude()+",
"+status.getGeoLocation().getLongitude()):user.getLocation())
                    + ") - " + tweet.getText();
}
} catch (Exception te) {
    te.printStackTrace(); System.out.println("Failed to search tweets:
" + te.getMessage());
    System.exit(-1);
}
return resultString;
}
```

```
public static void main(String[] args) {  
    GetTweetByLocation tt= new GetTweetByLocation();  
    String aa=tt.getSimpleTimeLine();  
    System.out.println(aa);  
}
```

# Output

@eatSheffield (Sheffield) - RT @barandgrillshef: #Sheffield if you had to order a cocktail what would it be, or would you just like a cup from @YorkshireTea ?@barandgrillshef (Leopold Square, Sheffield) - #Sheffield if you had to order a cocktail what would it be, or would you just like a cup from @YorkshireTea ?

@CFMDsFMKX (Sheffield Hallam University) - We're teaching today at #sheffieldhallam #sheffield on our UG programme in #facilitiesmanagement on Managing Premises & The Work Environment

@Map\_Game (-12.5743, 131.102) - Where is Sheffield on the map? Play the game at <http://www.map-game.com/sheffield> #Sheffield

@Map\_Game (-12.5743, 131.102) - Where is Sheffield on the map? Play the game at <http://www.map-game.com/sheffield> #Sheffield

@barandgrillshef (Leopold Square, Sheffield) - Fancy relaxing on the beach #sheffield <http://www.youtube.com/watch?v=Dax5Sbt20sA> we'll see you there

@barandgrillshef (Leopold Square, Sheffield) - #Sheffield #Cloudy according to the BBC <http://news.bbc.co.uk/weather/forecast/353> hows your day?

@barandgrillshef (Leopold Square, Sheffield) - #mothersday april 3 any plans #sheffield ? why not book a table now <http://www.barandgrillsheffield.co.uk/mothers-day/>

@Kineets (sheffield) - @shefgossip what's all the factor lot doing here @katiewaisel24 checked in #sheffield an hour ago?

@aryayuyutsu (53.382419,-1.478586) - RT @SheffieldStar 400 workers lose job as firm closes down in #Chesterfield <http://bit.ly/hpX8NK> (#Sheffield)

@Map\_Game (-12.5743, 131.102) - Where is Sheffield on the map? Play the game at <http://www.map-game.com/sheffield> #Sheffield

@Map\_Game (-12.5743, 131.102) - Where is Sheffield on the map? Play the game at <http://www.map-game.com/sheffield> #Sheffield

@Map\_Game (-12.5743, 131.102) - Where is Sheffield on the map? Play the game at <http://www.map-game.com/sheffield> #Sheffield

@aryayuyutsu (53.382419,-1.478586) - Off for the final night of a most ROTFL-ing and LOL-ing and LMAO-ing #ComedyFestival 2011. I voted for the amazing #Thunderbards! #Sheffield

@Map\_Game (-12.5743, 131.102) - Where is Sheffield on the map? Play the game at <http://www.map-game.com/sheffield> #Sheffield



# Main Classes: Status

- It represents one tweet

java.util.Date	<b>getCreatedAt()</b> Return the created_at
long	<b>getCurrentUserRetweetId()</b> Returns the authenticating user's retweet's id of this tweet, or -1L when the tweet is not a retweet.
GeoLocation	<b>getGeoLocation()</b> Returns The location that this tweet refers to if available.
long	<b>getId()</b> Returns the id of the status
java.lang.String	<b>getInReplyToScreenName()</b> Returns the in_reply_to_screen_name
long	<b>getInReplyToStatusId()</b> Returns the in_reply_tostatus_id
long	<b>getInReplyToUserId()</b> Returns the in_reply_user_id
Place	<b>getPlace()</b> Returns the place attached to this status
long	<b>getRetweetCount()</b> Returns the number of times this tweet has been retweeted, or -1 when the tweet is not a retweet.
Status	<b>getRetweetedStatus()</b>
java.lang.String	<b>getSource()</b> Returns the source
java.lang.String	<b>getText()</b> Returns the text of the status
User	<b>getUser()</b> Return the user associated with the status. This can be null if the instance is from User.getStatus().
boolean	<b>isFavorited()</b> Test if the status is favorited
boolean	<b>isPossiblySensitive()</b> Returns true if the status contains a link that is identified as sensitive.
boolean	<b>isRetweet()</b>

# Main Classes: User

- It represents a user

java.util.Date	<b>getCreatedAt()</b>
java.lang.String	<b>getDescription()</b> Returns the description of the user
URLEntity[]	<b>getDescriptionURLEntities()</b> Returns URL entities for user description.
int	<b>getFavouritesCount()</b>
int	<b>getFollowersCount()</b> Returns the number of followers
int	<b>getFriendsCount()</b>
long	<b>getId()</b> Returns the id of the user
java.lang.String	<b>getLang()</b> Returns the preferred language of the user
int	<b>getlistedCount()</b> Returns the number of public lists the user is listed on
java.lang.String	<b>getLocation()</b> Returns the location of the user
java.lang.String	<b>getMiniProfileImageURL()</b>
java.lang.String	<b>getMiniProfileImageURLHttps()</b>
java.lang.String	<b>getName()</b> Returns the name of the user
java.lang.String	<b>getOriginalProfileImageURL()</b>

# User (2)

java.lang.String	<b>getScreenName()</b> Returns the screen name of the user
<b>Status</b>	<b>getStatus()</b> Returns the current status of the user This can be null if the instance is from Status.getUser().
int	<b>getStatusesCount()</b>
java.lang.String	<b>getTimeZone()</b>
java.lang.String	<b>getURL()</b> Returns the url of the user
<b>URLEntity</b>	<b>getURLEntity()</b> Returns URL entity for user's URL.
int	<b>getUtcOffset()</b>
boolean	<b>isContributorsEnabled()</b> Tests if the user is enabling contributors
boolean	<b>isFollowRequestSent()</b> Returns true if the authenticating user has requested to follow the user.
boolean	<b>isGeoEnabled()</b>
boolean	<b>isProfileBackgroundTiled()</b>
boolean	<b>isProfileUseBackgroundImage()</b>
boolean	<b>isProtected()</b> Test if the user status is protected
boolean	<b>isShowAllInlineMedia()</b>
boolean	<b>isTranslator()</b>
boolean	<b>isVerified()</b>

# Main Classes: Twitter



The  
University  
Of  
Sheffield.

User	<code>createFriendship(long userId)</code> Allows the authenticating users to follow the user specified in the ID parameter. Returns the befriended user in the requested format when successful.
User	<code>createFriendship(long userId, boolean follow)</code> Allows the authenticating users to follow the user specified in the ID parameter. Returns the befriended user in the requested format when successful.
User	<code>createFriendship(java.lang.String screenName)</code> Allows the authenticating users to follow the user specified in the ID parameter. Returns the befriended user in the requested format when successful.
User	<code>createFriendship(java.lang.String screenName, boolean follow)</code> Allows the authenticating users to follow the user specified in the ID parameter. Returns the befriended user in the requested format when successful.
User	<code>destroyFriendship(long userId)</code> Allows the authenticating users to unfollow the user specified in the ID parameter. Returns the unfollowed user in the requested format when successful.
User	<code>destroyFriendship(java.lang.String screenName)</code> Allows the authenticating users to unfollow the user specified in the ID parameter. Returns the unfollowed user in the requested format when successful.
IDs	<code>getFollowersIDs(long cursor)</code> Returns an array of numeric IDs for every user the specified user is followed by.
IDs	<code>getFollowersIDs(long userId, long cursor)</code> Returns an array of numeric IDs for every user the specified user is followed by.
IDs	<code>getFollowersIDs(java.lang.String screenName, long cursor)</code> Returns an array of numeric IDs for every user the specified user is followed by.
PagableResponseList<User>	<code>getFollowersList(long userId, long cursor)</code> Returns a cursored collection of user objects for users following the specified user. At this time, results are ordered with the most recent following first — however, this ordering may introduce consistency issues.
PagableResponseList<User>	<code>getFollowersList(java.lang.String screenName, long cursor)</code> Returns a cursored collection of user objects for users following the specified user. At this time, results are ordered with the most recent following first — however, this ordering may introduce consistency issues.
IDs	<code>getFriendsIDs(long cursor)</code> Returns an array of numeric IDs for every user the authenticating user is following.
IDs	<code>getFriendsIDs(long userId, long cursor)</code> Returns an array of numeric IDs for every user the specified user is following.
IDs	<code>getFriendsIDs(java.lang.String screenName, long cursor)</code> Returns an array of numeric IDs for every user the specified user is following.
PagableResponseList<User>	<code>getFriendsList(long userId, long cursor)</code>

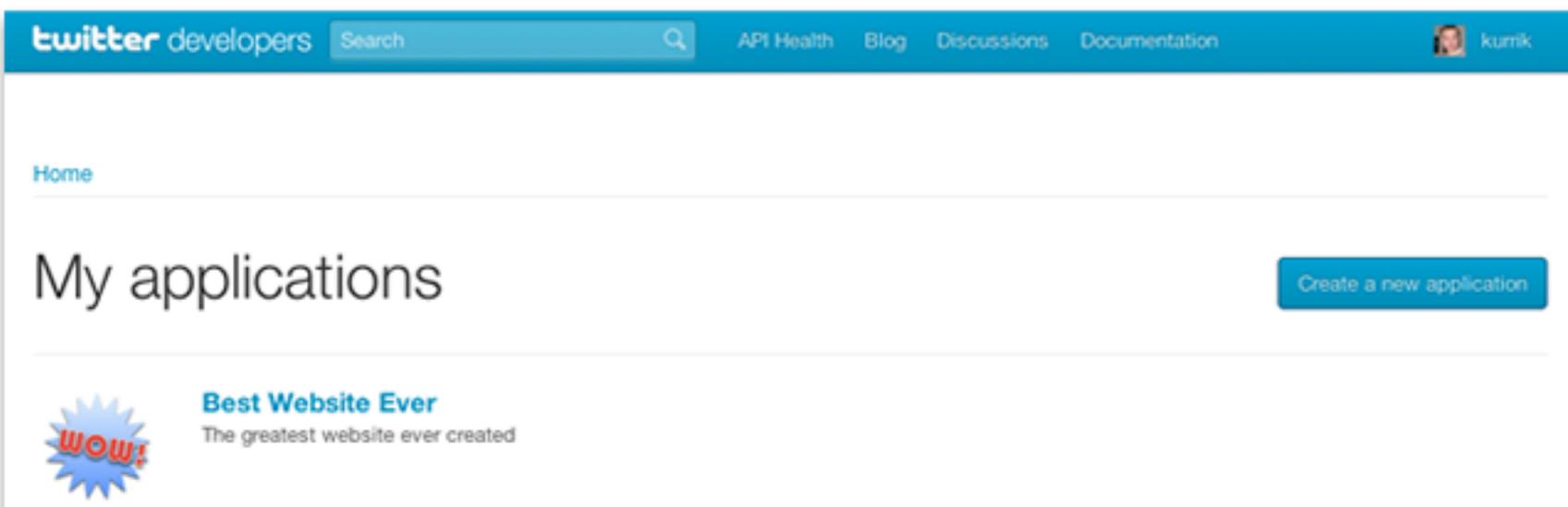
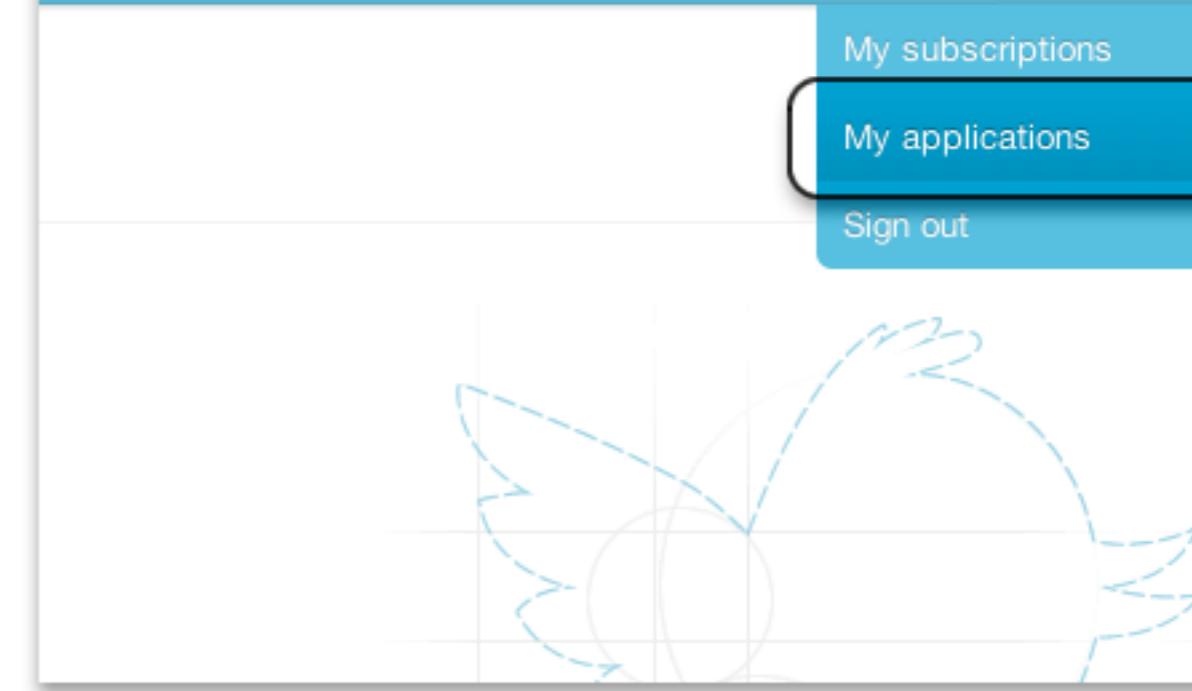
# Authentication for Twitter API

<https://dev.twitter.com/docs/auth/obtaining-access-tokens>

- In order to make authorized calls to Twitter's APIs
  - Your application must first obtain an OAuth access token
  - On behalf of a Twitter user
- The [dev.twitter.com](https://dev.twitter.com) application control panel offers the ability to generate an OAuth access token for the owner of the application.
  - This is useful if:
    - Your application only needs to make requests on behalf of a single user (for example, establishing a connection to the Streaming API)

# Generating a Token

- Visit [dev.twitter.com](https://dev.twitter.com/apps) "My applications" page, either by
  - navigating to [dev.twitter.com/apps](https://dev.twitter.com/apps),
  - or hovering over your profile image in the top right hand corner of the site and selecting "My applications"

A screenshot of the 'My applications' page on the Twitter developer portal. The page has a blue header with the 'twitter developers' logo, a search bar, and links for 'API Health', 'Blog', 'Discussions', and 'Documentation'. The user 'kurrik' is logged in. Below the header, there is a 'Home' link and a 'My applications' section. In this section, there is a card for an application named 'Best Website Ever' with the subtitle 'The greatest website ever created'. The card features a blue starburst icon with the word 'WOW!' in red. To the right of the card is a blue button labeled 'Create a new application'.

- Click on my applications  
--> Create new applications

# Access Token

- At the bottom of the next page, you will see a section labeled "your access token":

## Your access token

It looks like you haven't authorized this application for your own Twitter account yet. For your convenience, we give you the opportunity to create your OAuth access token here, so you can start signing your requests right away. The access token generated will reflect your application's current permission level.

[Create my access token](#)

- Click on the "Create my access token" button

## Your access token

Use the access token string as your "oauth\_token" and the access token secret as your "oauth\_token\_secret" to sign requests with your own Twitter account. Do not share your oauth\_token\_secret with anyone.

Access token

ddddddd-xxxxXxxxxXxxXXXXXxxxxXxxxxXxxxxXX

Access token secret

X00xx0x000xxxx000000xx0000xxxx000000xxxx

Access level

Read-only

[Recreate my access token](#)

```
1. OAuthTokens tokens = new OAuthTokens();  
2. tokens.ConsumerKey = "Consumer Key";  
3. tokens.ConsumerSecret = "Consumer Secret";  
4. tokens.AccessToken = "Access Key";  
5. tokens.AccessTokenSecret = "Access Secret";
```

- In Twitter4J

```
private Twitter initTwitter(String consumerKey, String consumerSecret,  
                           String accessToken, String accessTokenSecret) throws Exception {  
    ConfigurationBuilder cb = new ConfigurationBuilder();  
    cb.setDebugEnabled(true)  
        .setOAuthConsumerKey(consumerKey)  
        .setOAuthConsumerSecret(consumerSecret)  
        .setOAuthAccessToken(accessToken)  
        .setOAuthAccessTokenSecret(accessTokenSecret)  
        .setJSONStoreEnabled(true);  
    return (new TwitterFactory(cb.build()).getInstance());  
}
```

# Changing access level

- For most application the default access level (read-only) is fine
  - In some cases you will need writing permissions

trids USFD ← My Application Name

Details

Settings

OAuth tool

@Anywhere domains

Reset keys

Delete



monitoring twetts for emergency response  
[http://www.dcs.shef.ac.uk/~fabio/Fabio\\_Ciravegna/Projects.html](http://www.dcs.shef.ac.uk/~fabio/Fabio_Ciravegna/Projects.html)

Click settings

## Application Type

### Access:

- Read only
- Read and Write
- Read, Write and Access direct messages

What type of access does your application need? Note: @Anywhere applications require read & write access.

Find out more about our [Application Permission Model](#).

# LabClass

- Retrieving tweets for a specific location
- Displaying them on GoogleMaps

# Retrieving Friends (or Followers)

```
long[] tempFriendArray = new long[0];  
try {  
    long[] friendArray= twitter.getFriendsIDs(userId, -1).getIDs();  
    // followers: long[] followerArray= twitter.getFollowersIDs(userId, -1).getIDs();  
    Long[] myIds= new long[100]  
    For (int ix=0; ix<100; ix++) myIds[ix]= friendArray[ix];  
    ResponseList<twitter4j.User> userList = twitter.lookupUsers(myIds);  
    for (User us : ll) {  
        /* do whatever necessary with the user */  
    }  
} catch (TwitterException e) {  
    e.printStackTrace();  
}
```

It gets 5000 IDs at a time

It looks up up to 100 ids for one call

# Saving Twitter API calls

- Each OAuth key has 300 queries per hour allowed
- You always must check the code returned by each call
- If asked to desist you must stop and wait
  - Most calls will tell you when you can ask again
  - Sometimes they do not -> wait for an hour, then
- It is very important that you try to save calls to Twitter
- Be very careful during the assignment
  - Not listening to Twitter codes is unacceptable (you will lose marks)
  - Using multiple keys is forbidden
  - Not trying to save calls will lose you marks

# GoogleMaps

<http://maps.google.com/>

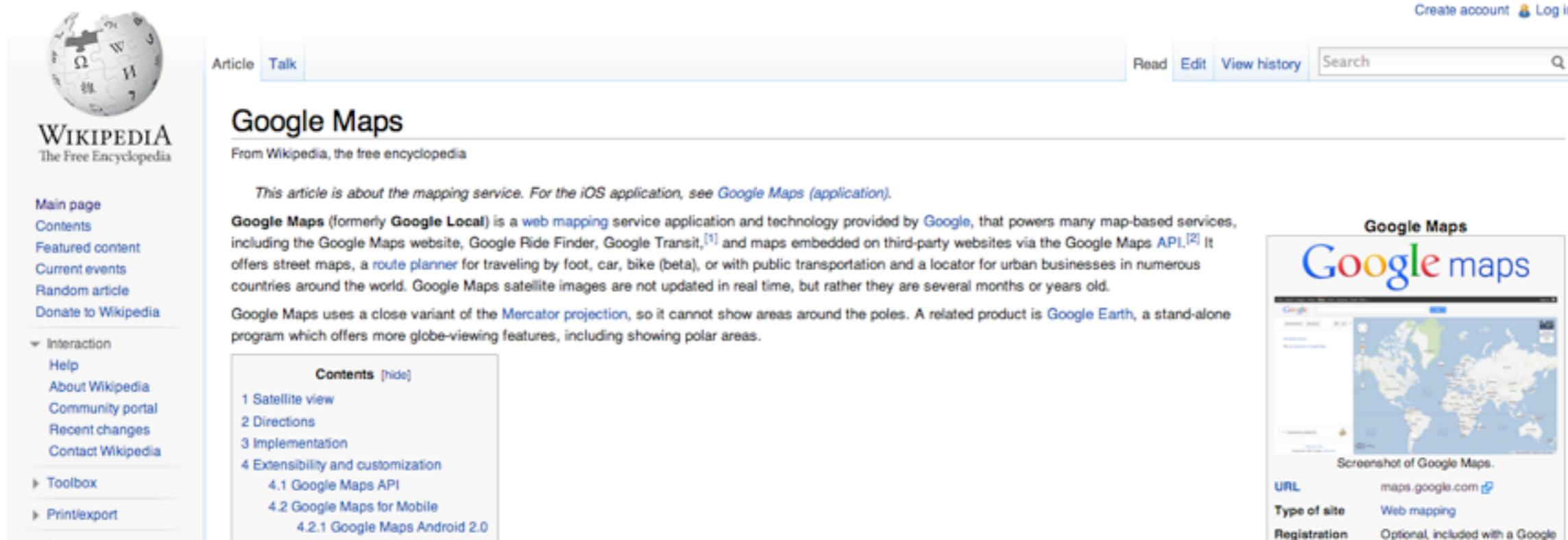


A screenshot of the Google Maps interface. At the top, there is a navigation bar with links for +You, Search, Images, Maps, Play, YouTube, News, Gmail, Drive, Calendar, and More. Below the bar, the word "Google" is written in its signature multi-colored font. To the right of "Google" is a search bar with a magnifying glass icon. Underneath the search bar are two buttons: "Get directions" and "My places". Further down are icons for printing and sharing, followed by a location pin and a compass rose. A sidebar on the left displays the current location as "Sheffield, South Yorkshire, UK" and includes a link to "Correct it ▾". It also features a section titled "Experience MapsGL" with a 3D building icon, listing four bullet points: "Take 3D photo tours of landmarks New!", "View 3D buildings", "Fly over 45-degree aerial view imagery", and "'Swoop' quickly into Street View, without a plugin". Below this, a note states "MapsGL is our Beta Maps technology powered by WebGL, and has certain system requirements." and a "Try it now" button. The main map area shows the British Isles with a dense network of roads. Major cities like London, Birmingham, Manchester, and Edinburgh are labeled. The map also includes labels for the North Sea, English Channel, and the Isle of Man. To the right, parts of the Netherlands and Belgium are visible. A scale bar at the bottom left indicates a distance of 100 mi.

# GoogleMaps

Article Talk Read Edit View history Search Q

Create account Log in



The screenshot shows the Wikipedia article page for "Google Maps". The page title is "Google Maps" and it is described as "From Wikipedia, the free encyclopedia". A note at the top states: "This article is about the mapping service. For the iOS application, see [Google Maps \(application\)](#)". The main content describes Google Maps as a web mapping service provided by Google, mentioning its street maps, route planner, and locator features. It notes that satellite images are not real-time. Below the main text is a sidebar with a "Contents" section listing topics like Satellite view, Directions, Implementation, and Extensibility and customization. To the right of the main content is a box titled "Google Maps" containing a screenshot of the Google Maps website showing a world map, and a summary table with details such as URL, Type of site, and Registration.

Google Maps	
	Screenshot of Google Maps.
URL	<a href="http://maps.google.com">maps.google.com</a>
Type of site	Web mapping
Registration	Optional, included with a Google

# What is relevant in GoogleMaps?

- Mapping-related web services had been available for some time from GIS vendors such as ESRI as well as from MapQuest and Microsoft MapPoint.
  - Experimenting with any of the formal vendor-supported web services required a formal contract between the parties
- Google Maps was implemented left the data for the taking, and hackers soon found ways to creatively re-use that data
  - Google has implemented an API for the Maps in a way similar to the one for the search engines
    - Remember that Google's idea was to provide tools for the community
    - This is also a generic trend in the new Web:
      - Design for "hackability" and remixability.
      - Systems like the original web, RSS, and AJAX all have this in common: the barriers to re-use are extremely low.
      - Much of the useful software is actually open source, but even when it isn't, there is little in the way of intellectual property protection.

# Google Maps Javascript API V3

Google Maps JavaScript API v3  1.6k

Developer's Guide

API Reference

Code Samples

More Resources

Google Maps API for  
Business

## V3: The Solution for Maps Applications for both the Desktop and Mobile Devices

Note: The Google Maps Javascript API Version 3 is now the official Javascript API. Version 2 of this API has been officially deprecated as per our deprecation policy. We encourage you to migrate your code to this newly updated and enhanced version!

The Google Maps Javascript API lets you embed Google Maps in your own web pages. Version 3 of this API is especially designed to be faster and more applicable to mobile devices, as well as traditional desktop browser applications.

The API provides a number of utilities for manipulating maps (just like on the <http://maps.google.com> web page) and adding content to the map through a variety of markers, overlays, and controls, allowing you to create robust maps applications on your website.

<http://code.google.com/apis/maps/documentation/javascript/tutorial.html>

# The Hello World of GoogleMaps

```
<!DOCTYPE html>
<html>
<head>
<meta name="viewport" content="initial-scale=1.0, user-scalable=no" />
<style type="text/css">
  html { height: 100% }
  body { height: 100%; margin: 0px; padding: 0px }
  #map_canvas { height: 100% }
</style>
<script type="text/javascript"
  src="http://maps.google.com/maps/api/js?sensor=set_to_true_or_false">
</script>
<script type="text/javascript">
  function initialize() {
    var latlng = new google.maps.LatLng(-34.397, 150.644);
    var myOptions = {  zoom: 8,
  center: latlng,
  mapTypeId: google.maps.MapTypeId.ROADMAP
    };
    var map = new google.maps.Map(document.getElementById("map_canvas"),
      myOptions);
  }
</script>
</head>
<body onload="initialize()">
  <div id="map_canvas" style="width:100%; height:100%"></div>
</body>
</html>
```

Maps API JavaScript using a **script** tag.

Called when page is loaded

JavaScript object literal to hold a number of map properties

ROADMAP or SATELLITE or HYBRID or TERRAIN

div element "map\_canvas" to hold the Map

# Map Class

## google.maps.Map - the Elementary Object

```
var map = new google.maps.Map(document.getElementById("map_canvas"),  
    myOptions);
```

The JavaScript class that represents a map is the `Map` class. Objects of this class define a single map on a page. (You may create more than one instance of this class - each object will define a separate map on the page.) We create a new instance of this class using the JavaScript `new` operator.

When you create a new map instance, you specify a `<div>` HTML element in the page as a container for the map. HTML nodes are children of the JavaScript `document` object, and we obtain a reference to this element via the `document.getElementById()` method.

This code defines a variable (named `map`) and assigns that variable to a new `Map` object, also passing in options defined within the `myOptions` object literal. These options will be used to initialize the map's properties. The function `Map()` is known as a *constructor* and its definition is shown below:

Constructor	Description
<code>google.maps.Map(opts? )</code>	Creates a new map using the passed optional parameters in the <code>opts</code> parameter.

# Geocoding

This department:

- Latitude is: 53.38108855193859
- Longitude is: -1.4801287651062012

## Geocoding Requests

A Geocoding API request must be of the following form:

```
http://maps.googleapis.com/maps/api/geocode/output?parameters
```

where `output` may be either of the following values:

- `json` (recommended) indicates output in JavaScript Object Notation (JSON)
- `xml` indicates output as XML

Certain parameters are required while some are optional. As is standard in URLs, all parameters are separated using the ampersand (`&`) character. The list of parameters and their possible values are enumerated below.

The Geocoding API defines a geocoding request using the following URL parameters:

- `address` (required) — The address that you want to geocode.\*
- OR
- `latlng` (required) — The textual latitude/longitude value for which you wish to obtain the closest, human-readable address.\*
  - `bounds` (optional) — The bounding box of the viewport within which to bias geocode results more prominently. (For more information see [Viewport Biasing](#) below.)
  - `region` (optional) — The region code, specified as a ccTLD ("top-level domain") two-character value. (For more information see [Region Biasing](#) below.)
  - `language` (optional) — The language in which to return results. See the [supported list of domain languages](#). Note that we often update supported languages so this list may not be exhaustive. If `language` is not supplied, the geocoder will attempt to use the native language of the domain from which the request is sent wherever possible.
  - `sensor` (required) — Indicates whether or not the geocoding request comes from a device with a location sensor. This value must be either `true` or `false`.

\* Note: You may pass either an `address` or a `latlng` to lookup. (If you pass a `latlng`, the geocoder performs what is known as a *reverse geocode*. See [Reverse Geocoding](#) for more information.)

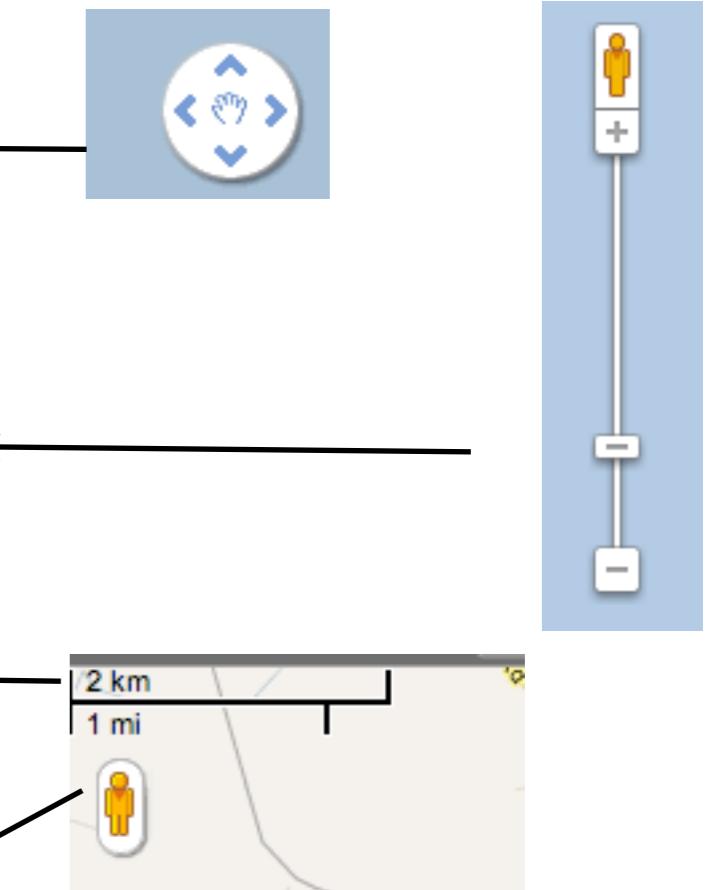
The `bounds` and `region` parameters will only influence, not fully restrict, results from the geocoder.

# Adding Control to the Map

```

function initialize() {
  var myOptions = {
    zoom: 12,
    center: new google.maps.LatLng(-28.643387, 153.612224),
    mapTypeId: google.maps.MapTypeId.ROADMAP,
    mapTypeControl: true,
    mapTypeControlOptions: {
      style: google.maps.MapTypeControlStyle.HORIZONTAL_BAR,
      position: google.maps.ControlPosition.BOTTOM_CENTER
    },
    panControl: true,
    panControlOptions: {
      position: google.maps.ControlPosition.TOP_RIGHT ←
    },
    zoomControl: true,
    zoomControlOptions: {
      style: google.maps.ZoomControlStyle.LARGE,
      position: google.maps.ControlPosition.LEFT_CENTER ←
    },
    scaleControl: true,
    scaleControlOptions: {
      position: google.maps.ControlPosition.TOP_LEFT ←
    },
    streetViewControl: true,
    streetViewControlOptions: {
      position: google.maps.ControlPosition.LEFT_TOP ←
    }
  }
  var map = new google.maps.Map(document.getElementById("map_canvas"),
    myOptions);
}

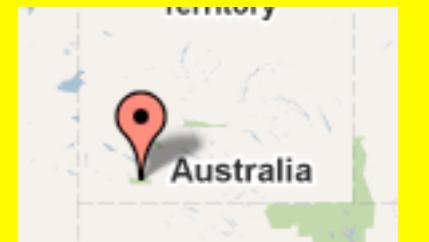
```



# Adding Markers and InfoWindows

```
var infowindow = new google.maps.InfoWindow({  
    content: "Hello!"  
});  
  
var marker = new google.maps.Marker({  
    position: myLatlng,  
    map: map,  
    title: "Uluru (Ayers Rock)"  
});  
  
google.maps.event.addListener(marker, 'click', function() {  
    infowindow.open(map, marker);  
});  
  
// To add the marker to the map, call setMap();  
marker.setMap(map);
```

Adds the marker



Adds the event associated to  
the marker

Uluru (Ayers Rock)  
Hello!

And Finally...

Or....

- Is it just another bubble?
  - Here Comes Another Bubble v1.1 - The Richter Scales
    - [http://www.youtube.com/watch?v=I6IQ\\_FOCE6I](http://www.youtube.com/watch?v=I6IQ_FOCE6I)
    - <http://www.richterscales.com/>
  - Winner of the Webby Award for Viral Video!



**"There's absolutely no bubble  
in technology"**

© Fabio Chavagna, University of Genoa