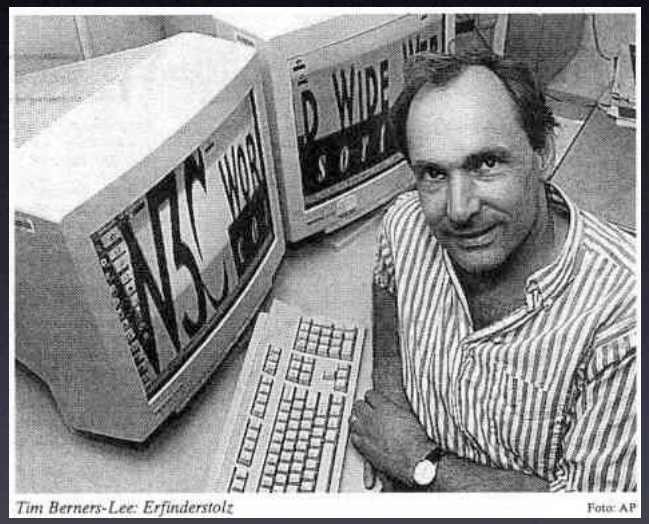
Sally Lynskey
Chief Executive

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Here Comes Everything: Preparing for the Linked Web of Data

Wendy Hall
Nigel Shadbolt

11 January 2010
Chilworth Manor



Tim Berners-Lee: Erfinderstolz Foto: AP

Vague but exciting ...

CERN DD/OC

Information Management: A Proposal

Tim Berners-Lee, CERN/DD

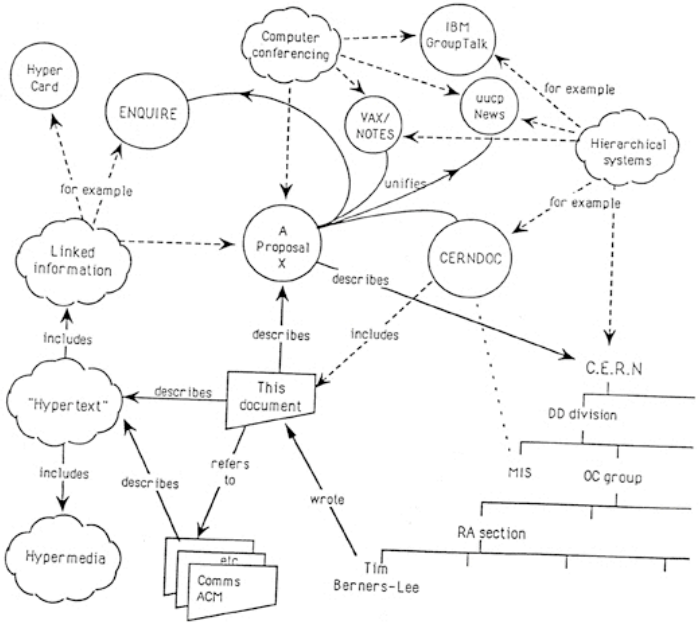
March 1989

Information Management: A Proposal

Abstract

This proposal concerns the management of general information about accelerators and experiments at CERN. It discusses the problems of loss of information about complex evolving systems and derives a solution based on a distributed hypertext sytstem.

Keywords: Hypertext, Computer conferencing, Document retrieval, Information management, Project control



Lessons learnt:

Big is beautiful:
the network is everything (tbl)

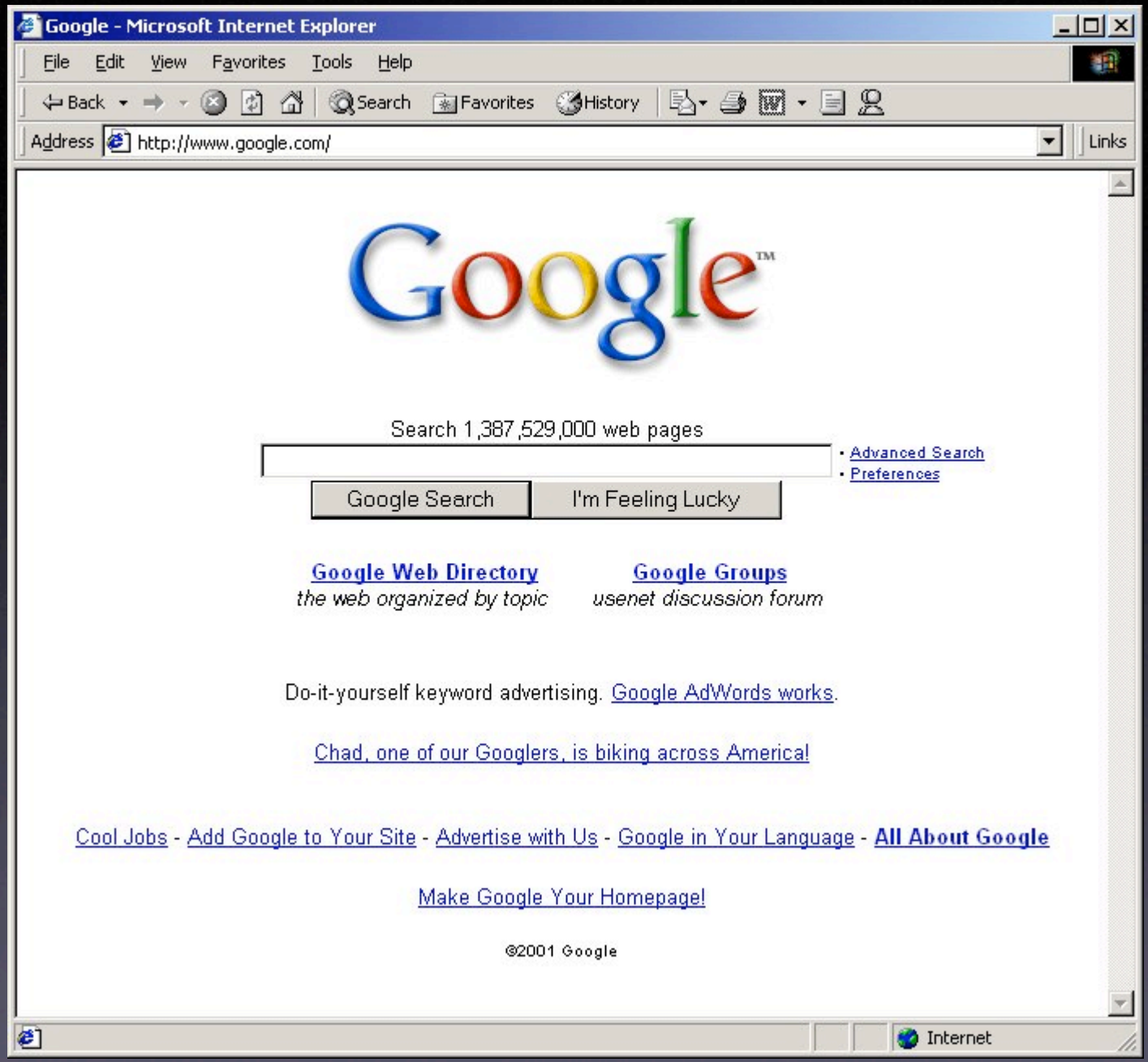
Scruffy works: let the links fail to make it scale (tbl)

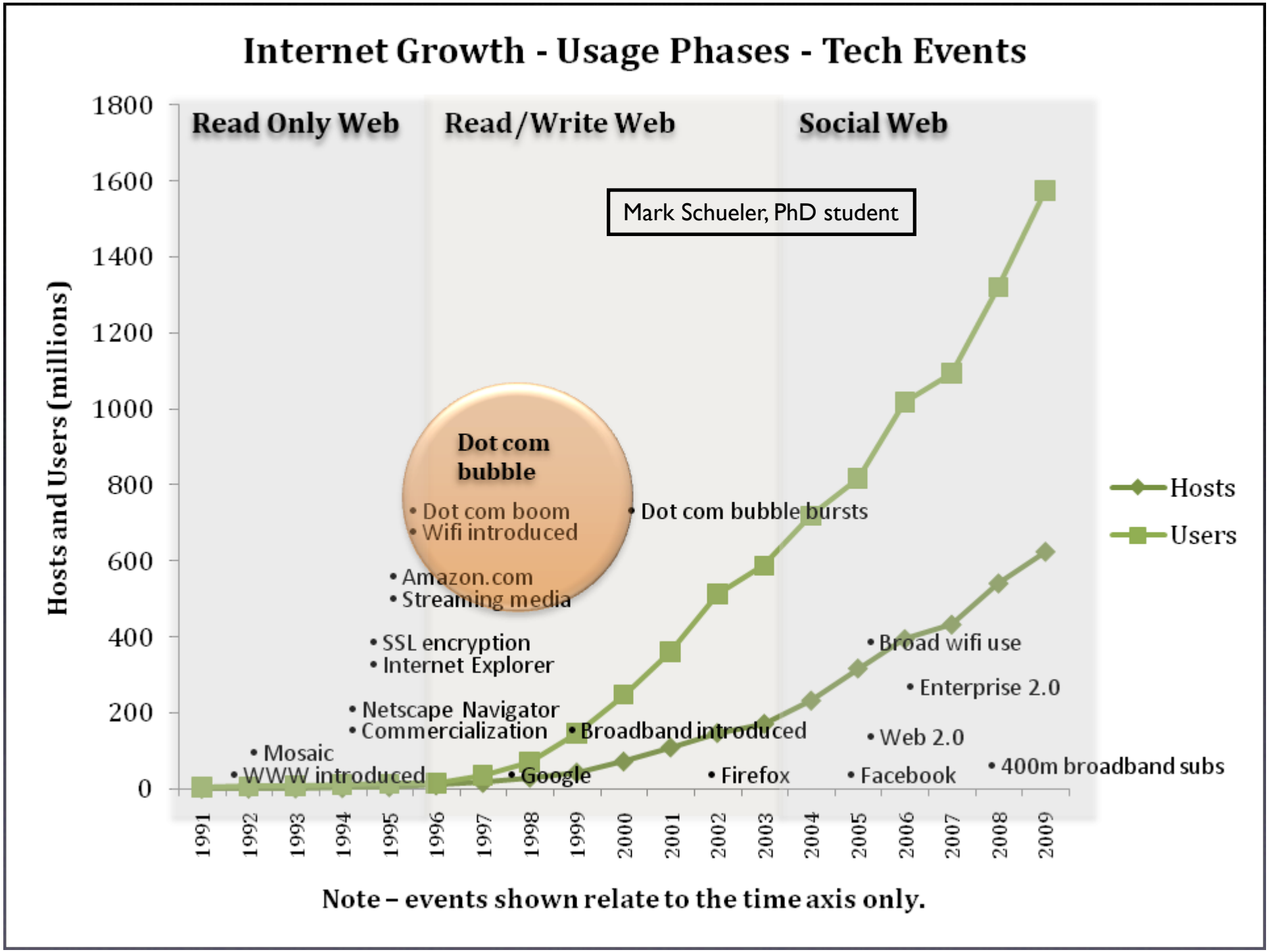
Democracy rules: open and free (tbl)

Open, non-proprietary, universal standards

But as the Web got bigger

Couldn’t find information easily just by following links


Google - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites History
Address http://www.google.com/
Links
Google™
Search 1,387,529,000 web pages
Advanced Search
Preferences
Google Search
I'm Feeling Lucky
Google Web Directory
the web organized by topic
Google Groups
usenet discussion forum
Do-it-yourself keyword advertising. Google AdWords works.
Chad, one of our Googlers, is biking across America!
Cool Jobs - Add Google to Your Site - Advertise with Us - Google in Your Language - All About Google
Make Google Your Homepage!
©2001 Google
Internet


Internet Growth - Usage Phases - Tech Events
Read Only Web
Read/Write Web
Social Web
Mark Schueler, PhD student
Hosts and Users (millions)
1800
1600
1400
1200
1000
800
600
400
200
0
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
Dot com bubble
• Dot com boom
• Wifi introduced
• Dot com bubble bursts
• Amazon.com
• Streaming media
• SSL encryption
• Internet Explorer
• Netscape Navigator
• Commercialization
• Broadband introduced
• Mosaic
• WWW introduced
• Google
• Firefox
• Facebook
• Broad wifi use
• Enterprise 2.0
• Web 2.0
• 400m broadband subs
Hosts
Users
Note – events shown relate to the time axis only.

Web 2.0

- Wiki's
- Blogs
- Folksonomies
- Flickr
- YouTube
- MySpace
- Facebook
- Second Life
- Twitter

Blogosphere

Wikipedia, the free encyclopedia - Windows Internet Explorer

http://en.wikipedia.org/wiki/Main_Page

File Edit View Favorites Tools Help

Wikipedia, the free ency... realtimecollisiondetection.ne...

Try Beta Log in / create account

main page | discussion | view source | history

WIKIPEDIA
The Free Encyclopedia

navigation
- Main page
- Contents
- Featured content
- Current events
- Random article

search
Go Search

interaction
- About Wikipedia
- Community portal
- Recent changes
- Contact Wikipedia
- Donate to Wikipedia
- Help

toolbox
- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link

[Hide]

Wikipedia is there when you need it — now it needs you.

$0.9M USD $7.5M USD

Donate Now

Welcome to Wikipedia,

the free encyclopedia that anyone can edit.

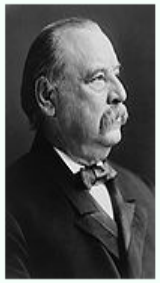
3,102,741 articles in English

- Arts
- Biography
- Geography
- History
- Mathematics
- Science
- Society
- Technology
- All portals

Overview · Editing · Questions · Help

Contents · Categories · Featured content · A–Z index

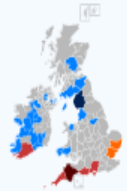
Today's featured article



Grover Cleveland (1837–1908) was both the 22nd and 24th President of the United States. Cleveland is the only President to serve two non-consecutive terms (1885–1889 and 1893–1897) and therefore is the only individual to be counted twice in the numbering of the presidents. He was the winner of the popular vote for President three times—in 1884, 1888, and 1892—and was the only Democrat elected to the Presidency in the era of Republican political domination that lasted from 1860 to 1912. Cleveland's admirers praise him for his honesty, independence, integrity, and commitment to the principles of classical liberalism. As a leader of the Bourbon Democrats, he opposed imperialism, taxes, subsidies and inflationary policies, but as a reformer he also worked against corruption, patronage, and bossism. Critics complained that he had little imagination and seemed overwhelmed by the nation's economic disasters—depressions and strikes—in his second term. Even so, his reputation for honesty and good character survived the troubles of his second term. (more...)

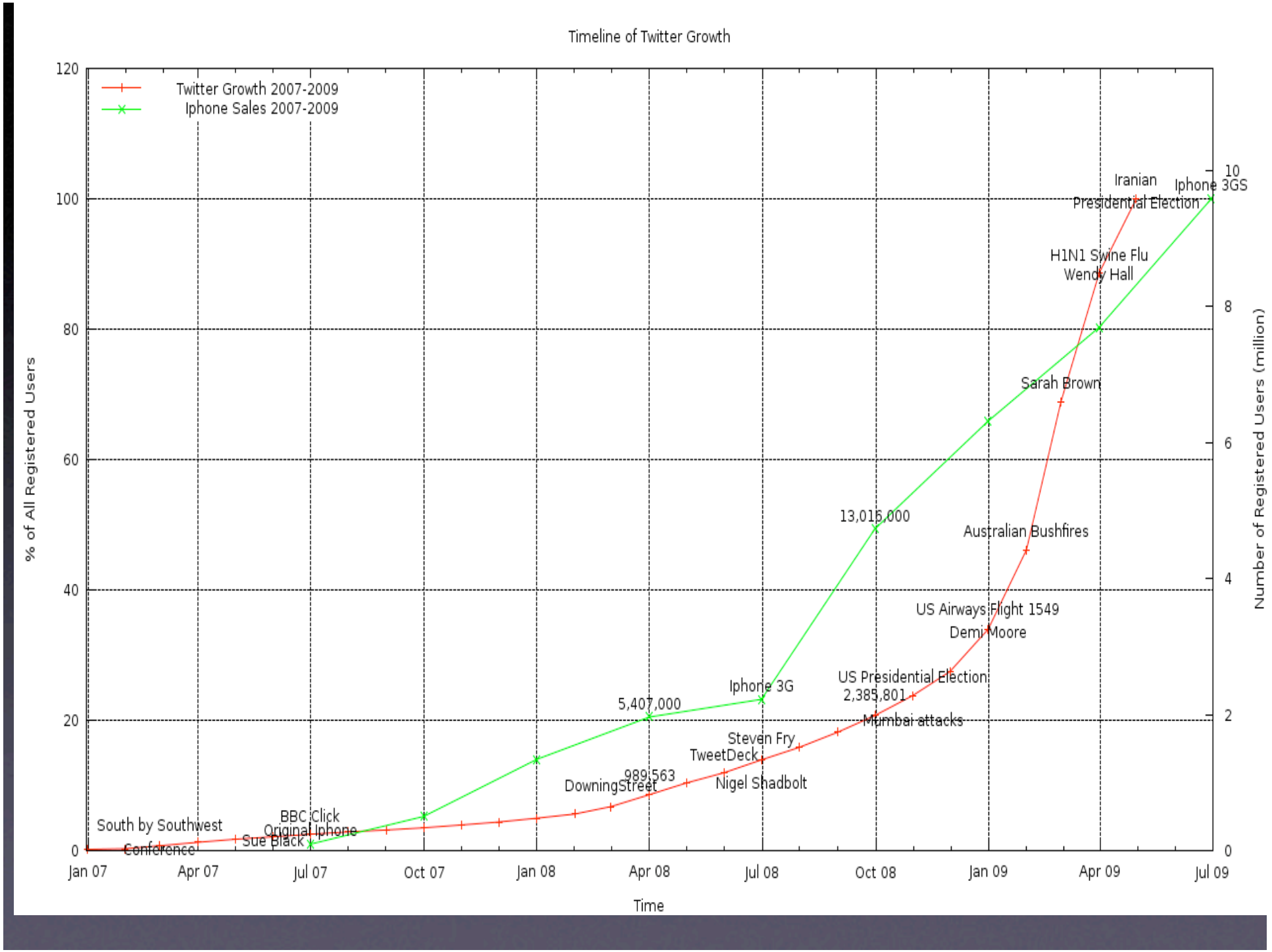
Recently featured: Werner Mölders – Electron – Han Dynasty

In the news



- At least 104 people are killed in a **mine explosion** in Heilongjiang, People's Republic of China.
- **Continued flooding** affecting Great Britain and Ireland *(affected areas pictured)* results in two deaths.
- Egypt **recalls** its ambassador to Algeria in a dispute following a qualifying match between the nations for the 2010 FIFA World Cup.
- Prime Minister of Belgium **Herman Van Rompuy** is chosen as the first permanent President of the European Council.
- British politician **Catherine Ashton** is selected as the first High Representative of the European Union for Foreign Affairs and Security Policy.
- A United States district court **finds** the Army Corps of Engineers was negligent in maintaining New Orleans levees that contributed to flooding during Hurricane Katrina.

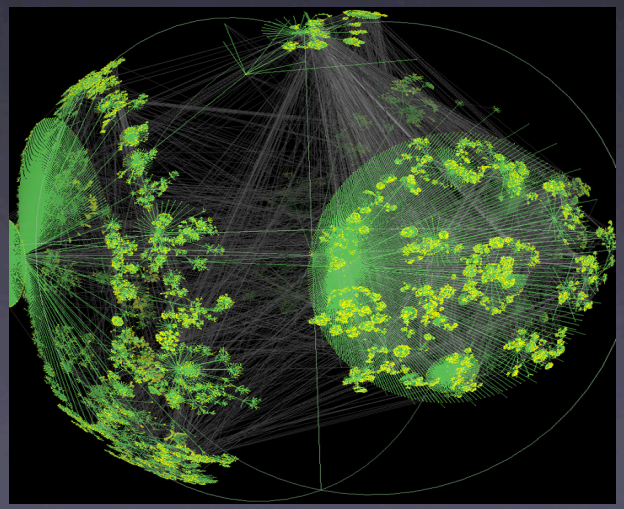
Internet 100%


Timeline of Twitter Growth
Twitter Growth 2007-2009
Iphone Sales 2007-2009
% of All Registered Users
Number of Registered Users (million)
Time
Jan 07
Apr 07
Jul 07
Oct 07
Jan 08
Apr 08
Jul 08
Oct 08
Jan 09
Apr 09
Jul 09
South by Southwest
Conference
Sue Black
BBC iClick
Original Iphone
5,407,000
Iphone 3G
13,016,000
DowningStreet
989,563
Nigel Shadbolt
TweetDeck
Steven Fry
US Presidential Election
2,385,801
Mumbai attacks
US Airways Flight 1549
Demi Moore
Australian Bushfires
Sarah Brown
H1N1 Swine Flu
Wendy Hall
Iranian
Presidential Election
Iphone 3GS

Galaxy Zoo 2 - Windows Internet Explorer
http://www.galaxyzoo.org/
Google
File Edit View Favorites Tools Help
Galaxy Zoo
Search
Sidewiki
Bookmarks
Check
Translate
AutoFill
Galaxy
Zoo
Sign In
Galaxy Zoo 2
Page
Tools
GALAXY ZOO 2
GALAXY ZOO
UNDERSTANDING COSMIC MERGERS
By colliding galaxies on your computer, you can help us trace the past, present and future of Galaxy Zoo mergers
How Do Galaxies Merge?
HELP US SOLVE THE PUZZLE
Home
The Story So Far
The Science
How To Take Part
Classify Galaxies
Forum
Zoo Media
Blog
FAQ
Contact Us
Welcome to Galaxy Zoo, where you can help astronomers explore the Universe
New, more detailed images added - see here for details
The Galaxy Zoo files contain almost a quarter of a million galaxies which
Classifier Log In
Click here to log in
Register
Forgotten Password?
Done
Internet
100%

Web 1.0

A Web of Documents



Web 2.0

- Wiki's
- Blogs
- Folksonomies
- Flickr
- YouTube
- MySpace
- Facebook
- Second Life
- Twitter

Blogosphere

Tomorrow the Web of Linked Data

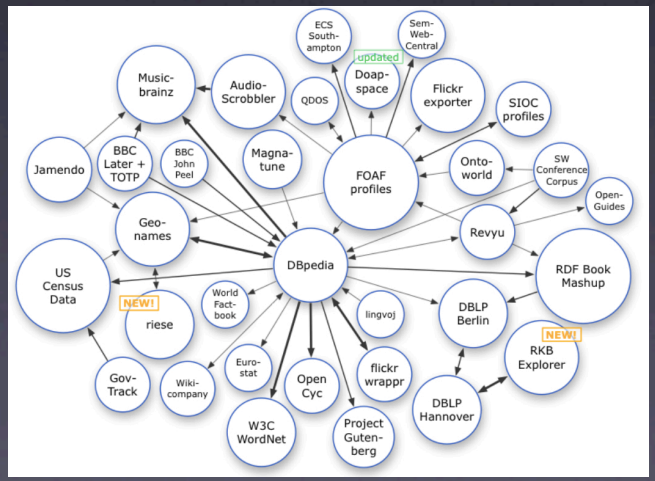
We're eager to share data for all sorts of reasons

What are the implications for business?

What are social and policy implications?

What will we do with it?

Web 3.0?



UNIVERSITY OF Southampton
School of Electronics and Computer Science

Linking Data

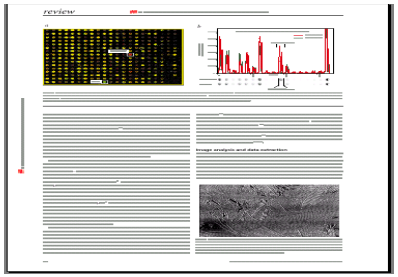
In a way that machines can "understand"....

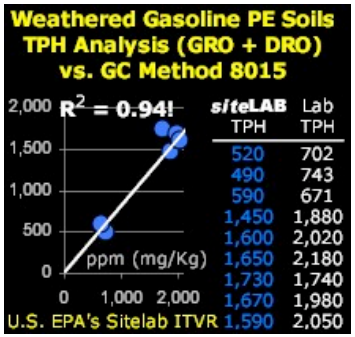
UNIVERSITY OF Southampton
School of Electronics and Computer Science

UNIVERSITY OF Southampton
School of Electronics and Computer Science

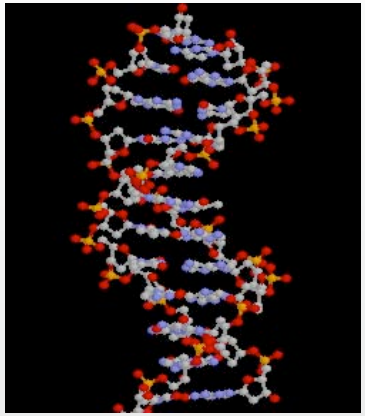
Can mark up and link any resource…

- Publications…
- Multimedia…
- Databases…
- Scientific structures…
- Workflow…
- People…





Web data set (XHTML)



UNIVERSITY OF Southampton
School of Electronics and Computer Science

Technical Principles

The four micro principles of the Semantic Web

1. All entities of interest, such as information resources, real-world objects, and vocabulary terms should be identified by URI references.
2. URI references should be dereferenceable, meaning that an application can look up a URI over the HTTP protocol and retrieve RDF data about the identified resource.
3. Data should be provided using the RDF/XML syntax
4. Data should be interlinked with other data.

The Semantic Web Revisited

Nigel Shadbolt and Wendy Hall, *University of Southampton*
Tim Berners-Lee, *Massachusetts Institute of Technology*

Nigel Shadbolt is a professor of artificial intelligence in the School of Electronics and Computer Science at Southampton University. Contact him at nrs@ecs.soton.ac.uk.

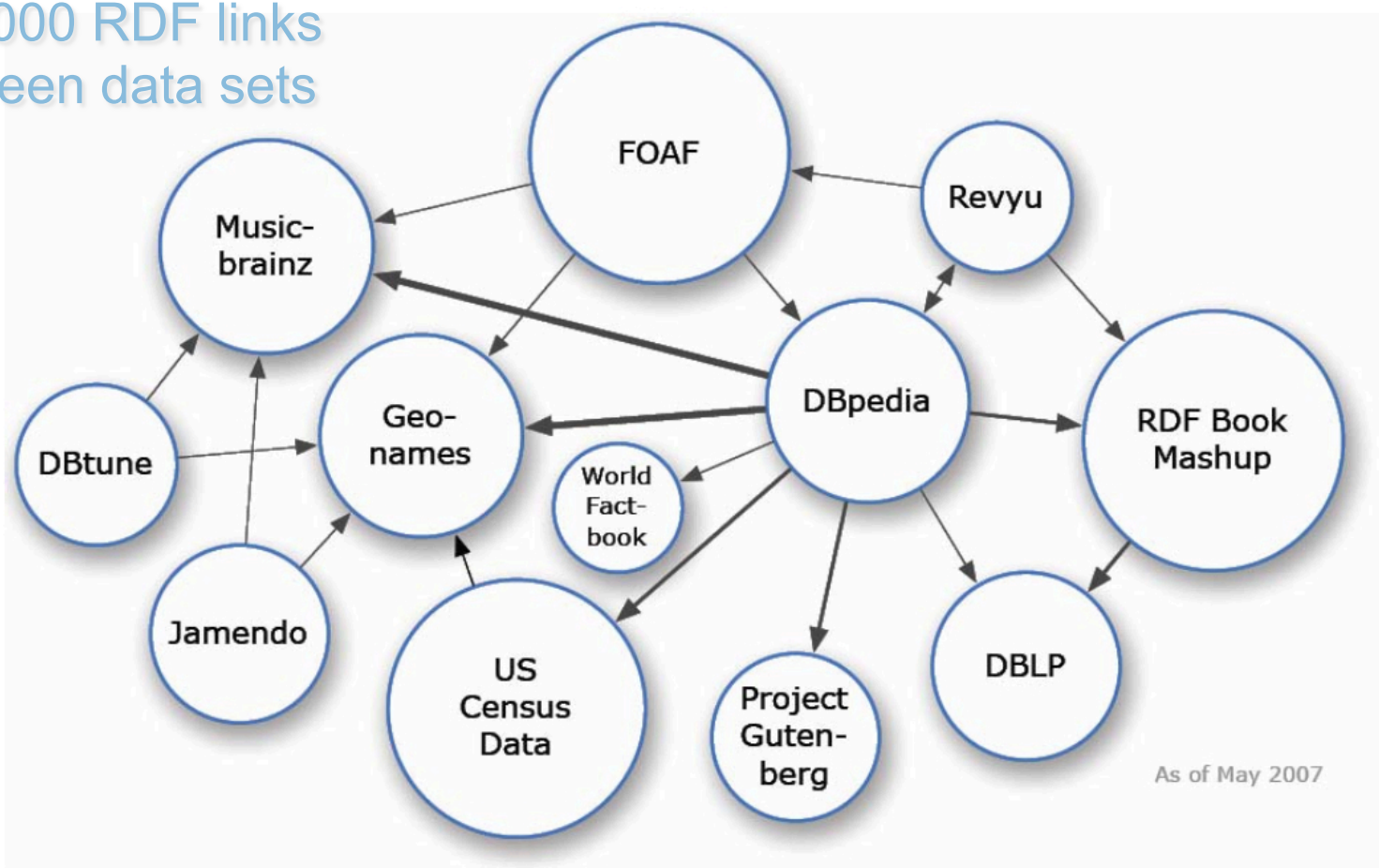
Tim Berners-Lee is the director of the World Wide Web Consortium, a senior researcher at the Massachusetts Institute of Technology's Computer Science and Artificial Intelligence Laboratory, and a professor of computer science in the Department of Electronics and Computer Science at Southampton University. Contact him at timbl@w3.org.

Wendy Hall is a professor of computer science in the School of Electronics and Computer Science at Southampton University. Contact her at wh@ecs.soton.ac.uk.

Linked Data on the Web: May 2007

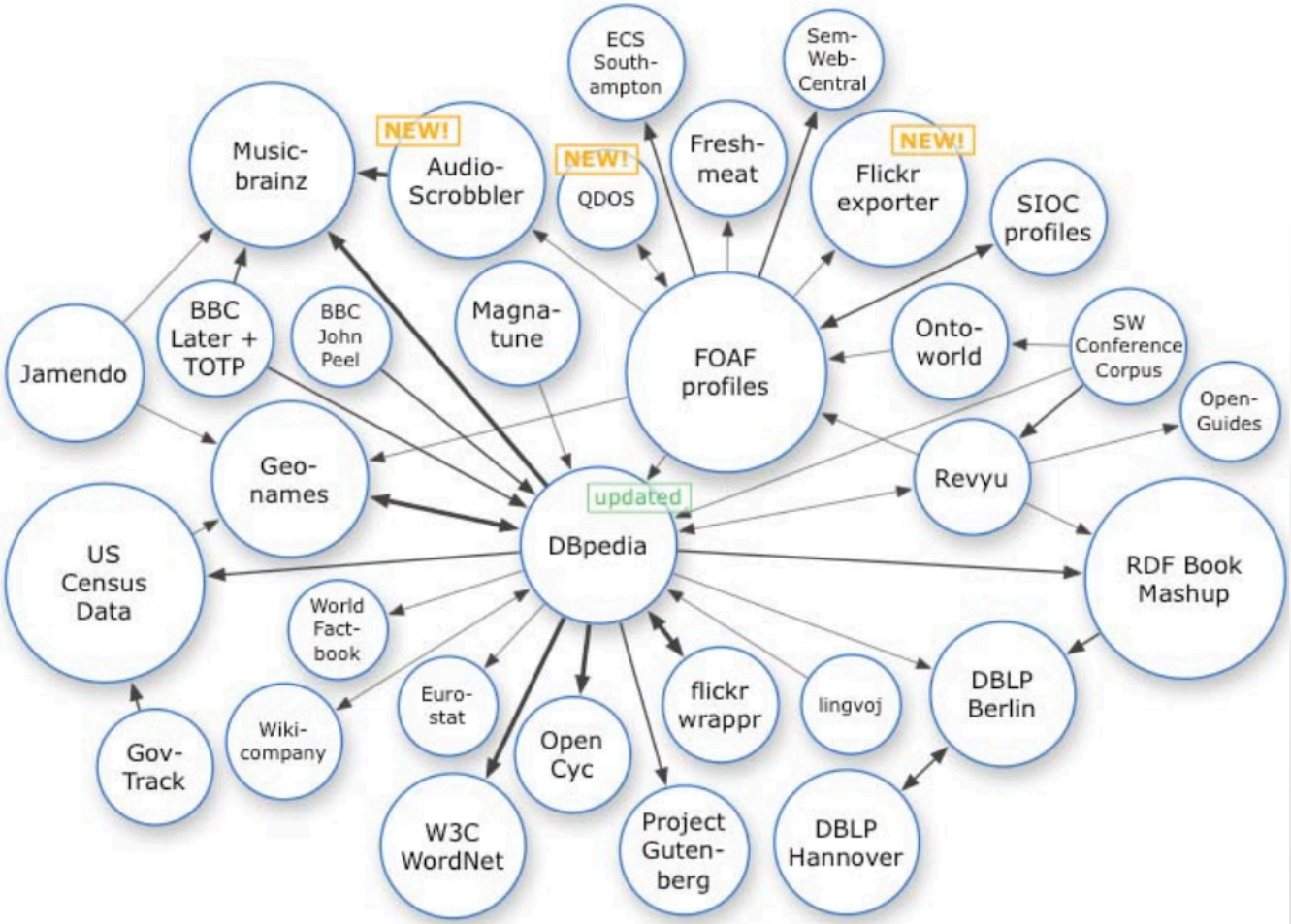
UNIVERSITY OF Southampton
School of Electronics and Computer Science

500 Million RDF Triples
120,000 RDF links
between data sets



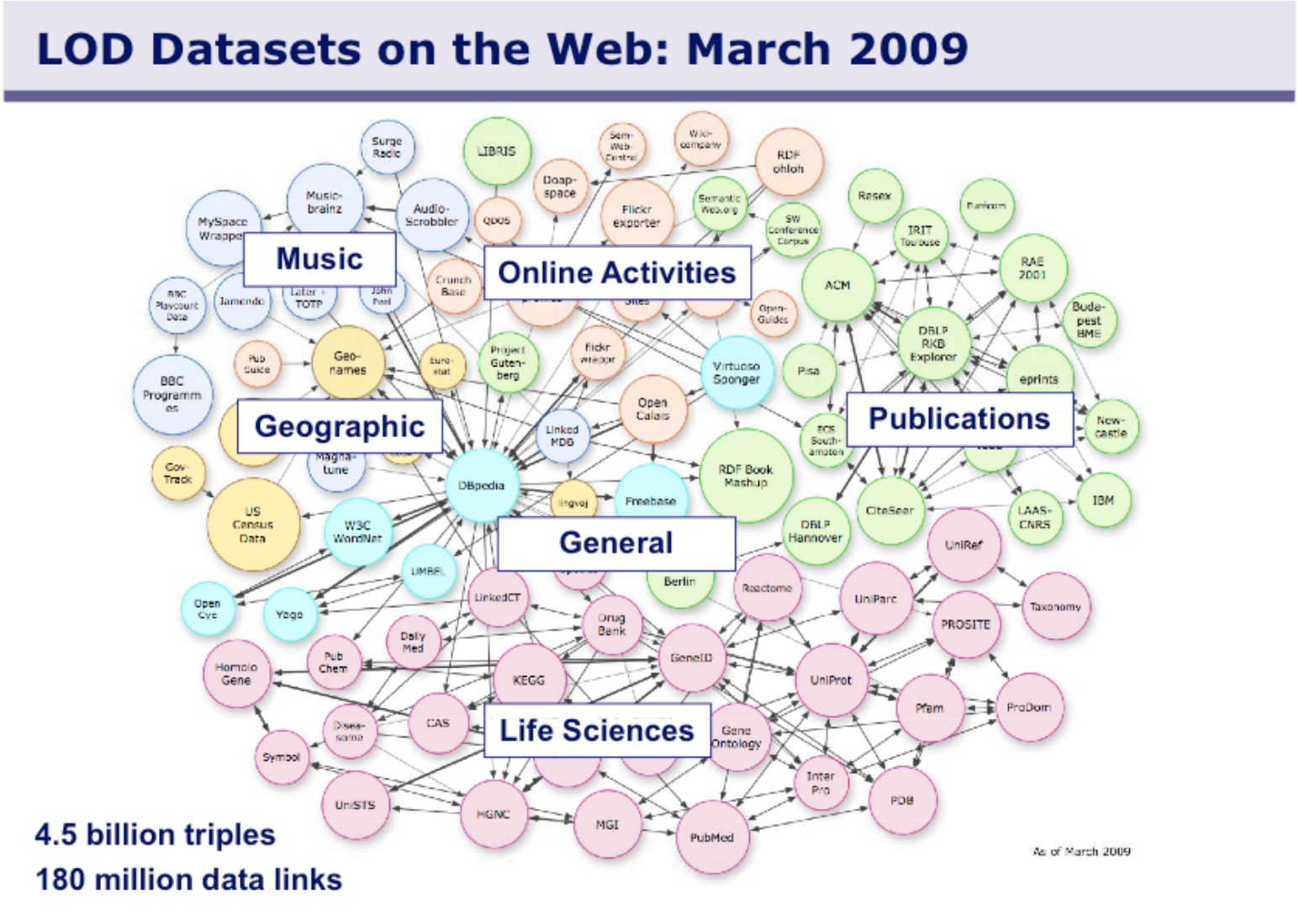
Linked Data on the Web: February 2008

UNIVERSITY OF Southampton
School of Electronics and Computer Science



Linked Data on the Web: March 2009

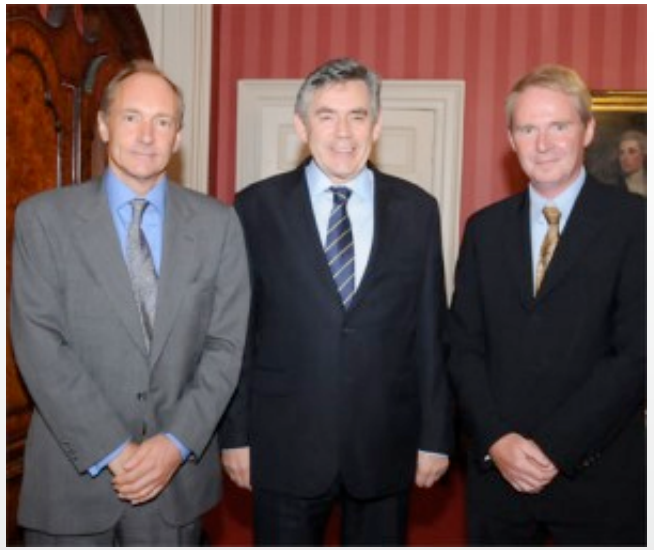
UNIVERSITY OF Southampton
School of Electronics and Computer Science



Linked Data in Government: A Tipping Point?

UNIVERSITY OF Southampton
School of Electronics and Computer Science

- Berners-Lee and Shadbolt - HMG Information Advisors
- Context - PoI Taskforce
- Previous work with OPSI
- Strategic Drivers
  - Transparency and accountability
  - Economic and Social Value
  - Public Service Improvement
  - New Industries New Jobs



UNIVERSITY OF Southampton
School of Electronics and Computer Science

Key Elemetns of our ToR

1. Establish single point of access for all public UK datasets - data.gov.uk
2. Specific proposals to implement and extend to wider public sector PSI principle
   - Select and implement common standards for release of public data
   - Select, develop and implement common terms for data where necessary
   - Develop licenses - CC variants
   - Support exploitation and publication of distributed and decentralised information assets
   - Potential for reform of info regulatory framework

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Progess to date

- Top down
  - Ministerial
  - CIO/CTOs
  - Engaged key data holders
- Middle Out
  - Civil Servants
  - Inventory of Information Assets
- Bottom up
  - Engaging with developer community
  - Cross Gov URI workshops
  - data.gov.uk
- Dissemination

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Progess to date

- Top down
  - Ministerial
  - CIO/CTOs
  - Engaged key data holders
- Middle Out
  - Civil Servants
  - Inventory of Information Assets
- Bottom up
  - Engaging with developer community
  - Cross Gov URI workshops
  - data.gov.uk
- Dissemination

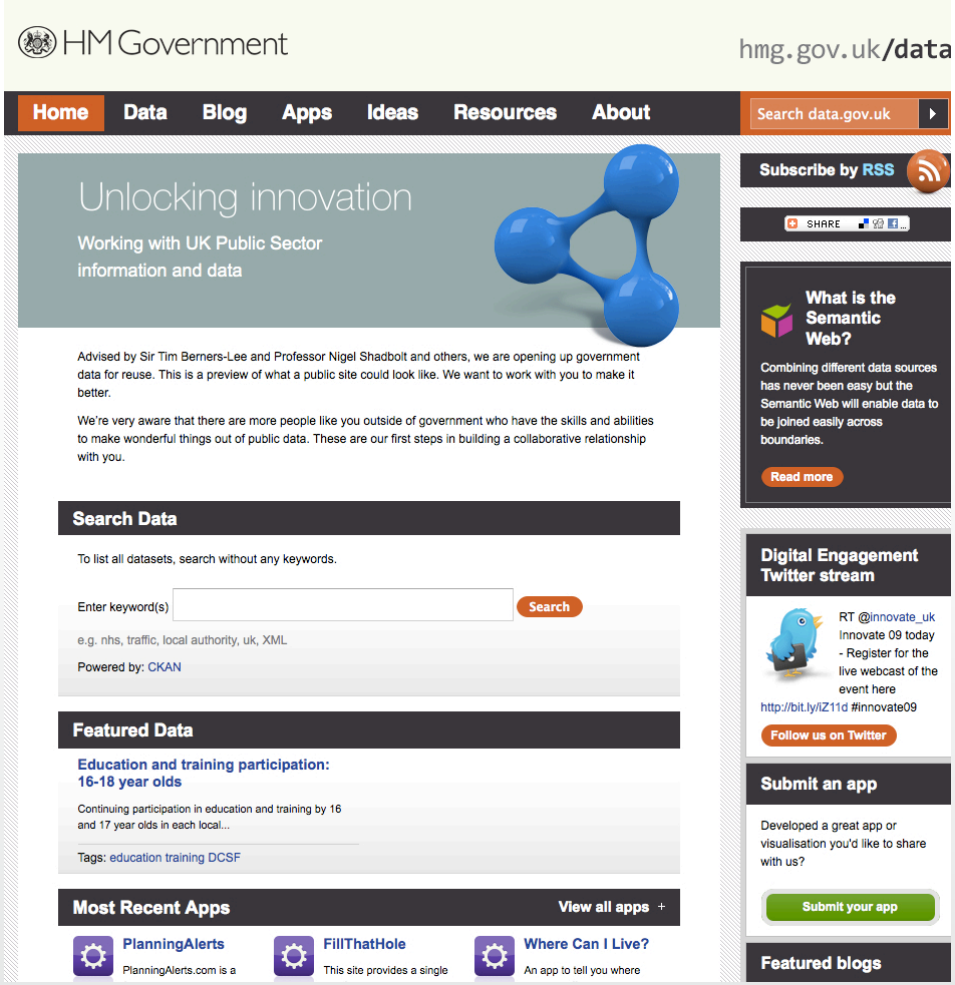
UNIVERSITY OF Southampton
School of Electronics and Computer Science

data.gov.uk

- Currently under development scrutiny and extension
  - Over 1500 datasets
  - Over 2000 registered developers
  - The beginning of linked data
- Beta service to be released 21st January 2010
- Access to OS from April 2010
- Clear schedule of release from Gov Depts thereafter

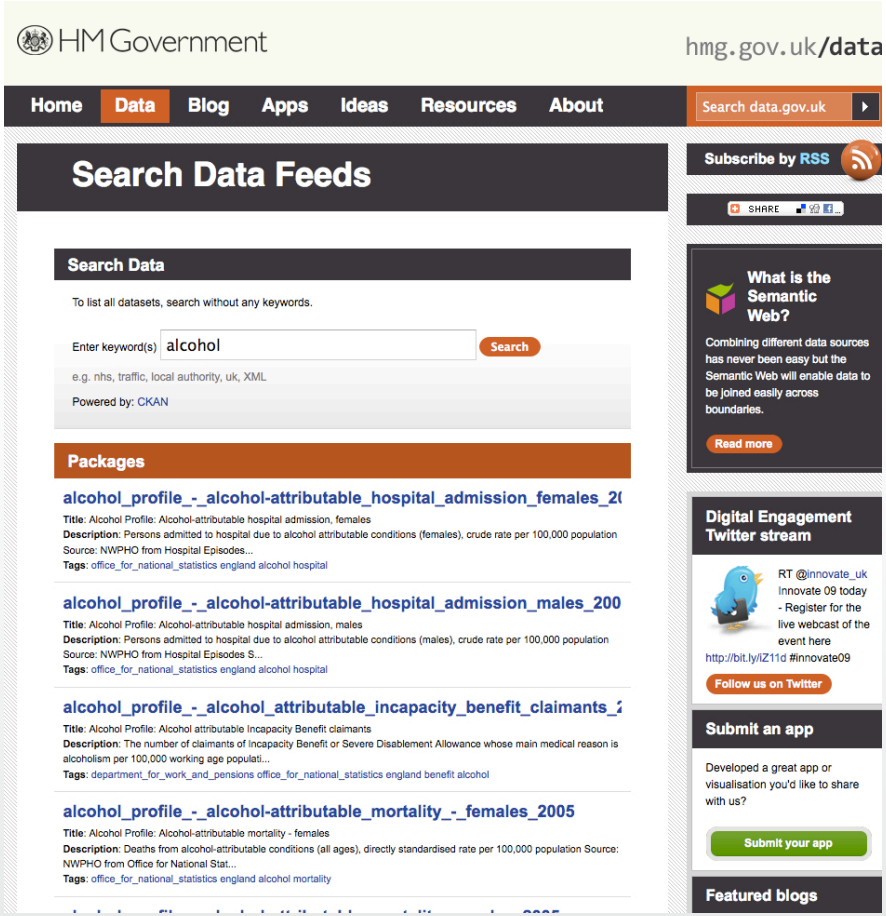
UNIVERSITY OF Southampton
School of Electronics and Computer Science

data.gov.uk



UNIVERSITY OF Southampton
School of Electronics and Computer Science

data.gov.uk



Alastair Dant (left) / David McCandless (right)

£ Massive Outgoings

Cost for average British taxpayer per day

GOVERNMENTS

ESSENTIALS

NHS £9.05

Schools £2.56

Defence £3.99

TRAVEL

Roads £3.10

LRT

£0.14

BANKING CHARGES

Recapitalisations

The Royal Bank of Scotland
Lloyds TSB
northern rock
Bradford & Bingley*

£9.23

SECURITY

Police

£0.53

Prisons

£0.19

Afghanistan

£0.32

GOING OUT

Arts

£0.04

Museums

£0.03

HELPING OTHERS

Child support

£1.86

Job Seekers

£0.25

Housing benefit

£1.52

Disability benefit

£1.44

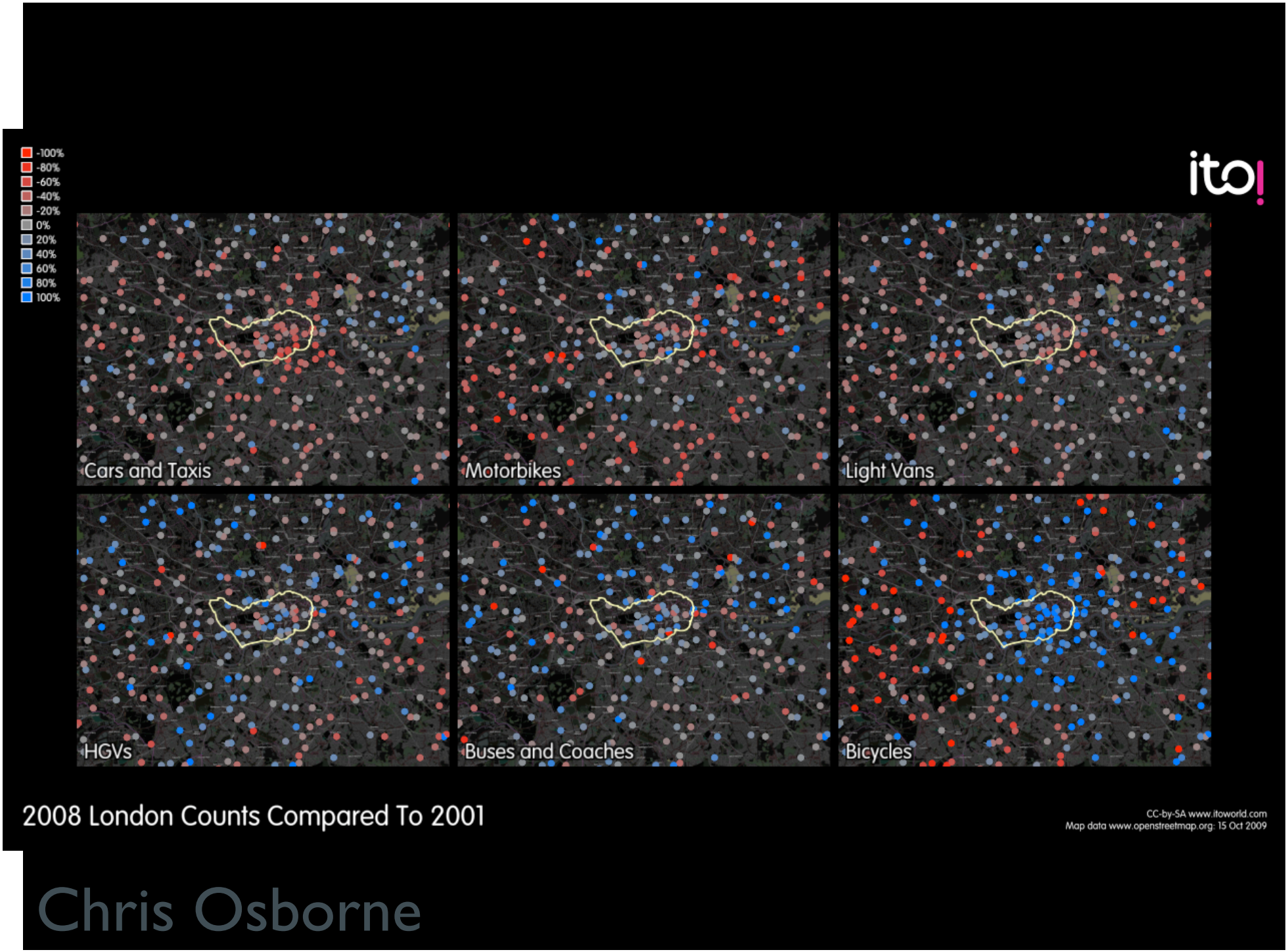
Pensions

£5.55

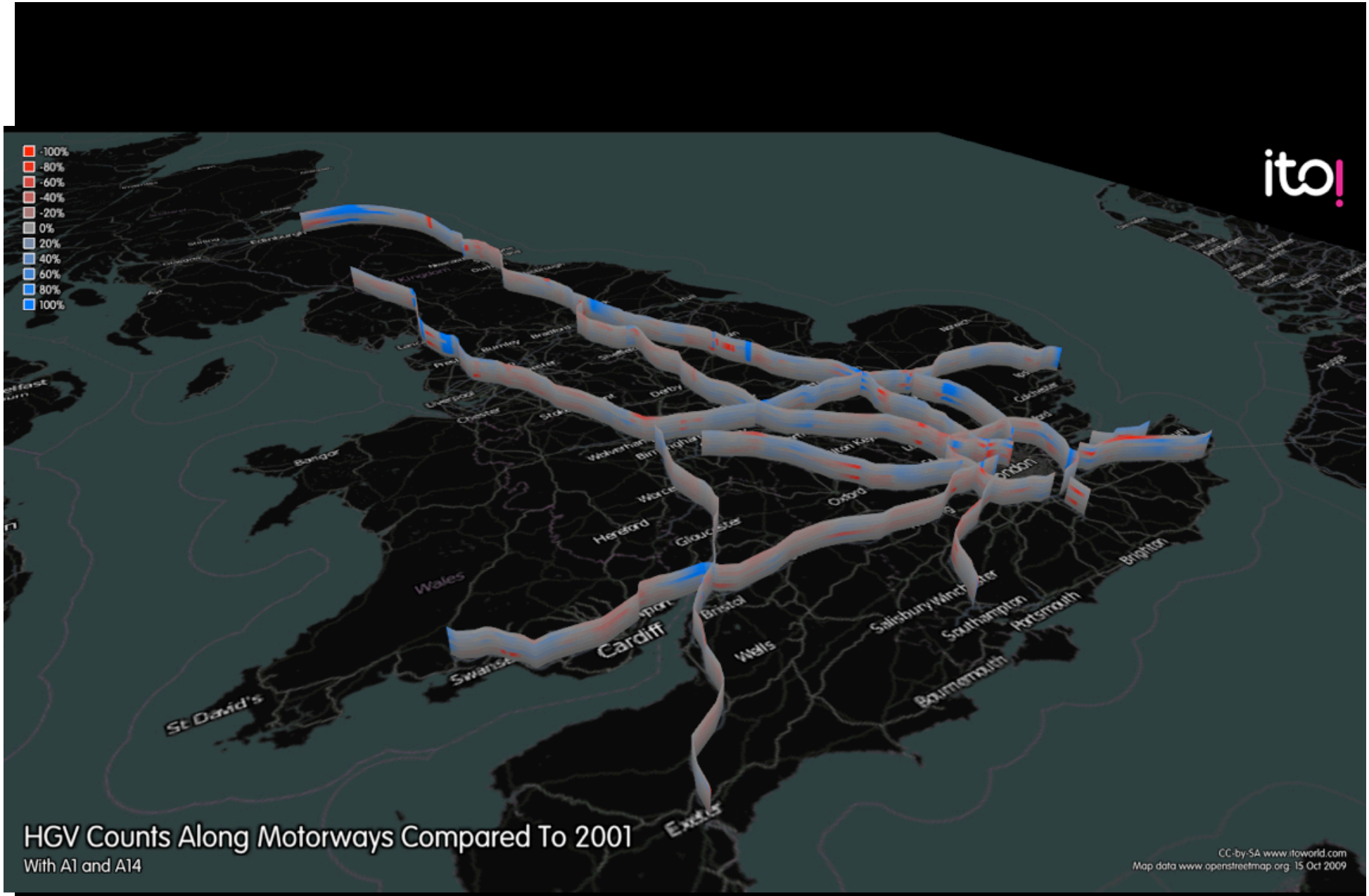
Overseas aid

£0.43

David McCandless

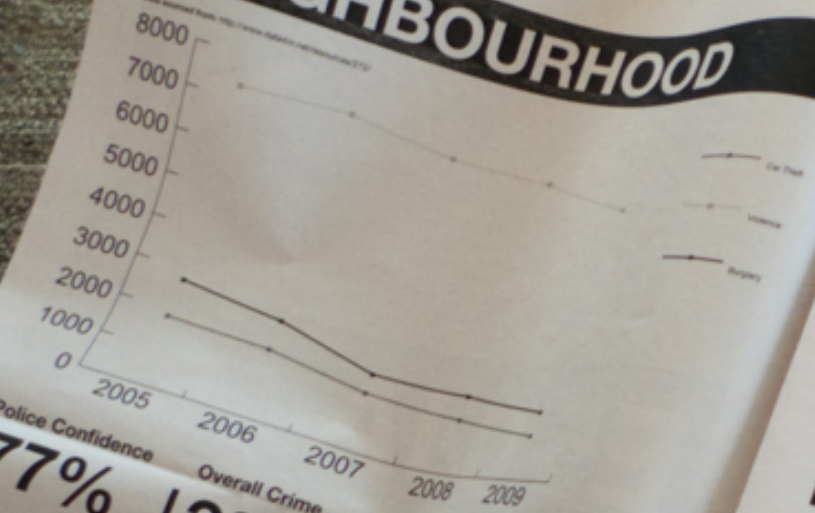


Chris Osborne



Chris Osborne

NEIGHBOURHOOD



Police Confidence
77%

Overall Crime
↓39%

Gun Crime
↓37%

Knife Crime
↓17%

Robbery
↓14%

Theft
↓13%

Woundings
↑220%

10000
fewer crime victims

Summary

Hackney Police and the Council work very closely to tackle and prevent crime in the borough. Seventy seven per cent of residents now have confidence in Hackney's Police, according to the British Crime Survey, a rise of 55 per cent in less than two years.

The survey also reveals that crime in Hackney has dropped 39.5 per cent over five years, the second largest drop of all London boroughs. Last year, there were ten thousand fewer victims of crime in Hackney than five years ago.

In 2008/09, gun crime was reduced by over 37 per cent and knife crime by 17 per cent. Youth crime also fell by 10 per cent. Satisfaction with the Police and the Fire Brigade are six per cent above the London average.

-- Hackney Council

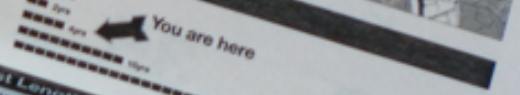
ALLOTMENTS

Allotment Plots

Tower Hamlet	282
Waltham Forest	1,382
Newham	415
Hackney	113
Haringey	1,715
Camden	195

Waiting List Time

Tower Hamlet
Waltham Forest
Newham
Hackney
Haringey
Camden



Waiting List Length

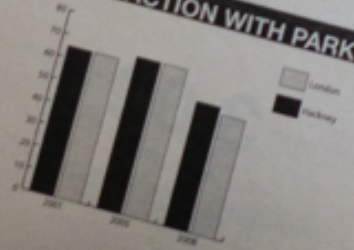
Tower Hamlet
Waltham Forest
Newham
Hackney
Haringey
Camden

GREEN SPACE

23%
by land area in Hackney

Your nearest parks are:
- Hackney Marshes (10 mins east)
- Victoria Park (15 mins south)
- London Fields (20 mins south west)

% SATISFACTION WITH PARKS



RECYCLING

Your nearest Recycling and Reuse Center is on South Access Road, Walthamstow, E17 (via Marshouse Avenue, near Low Hall Manor Business Centre).

Open every day of the year, except on 25 December, 26 December and 1 January. Open from 8.30am to 5.15pm. South Access Road is for mixed, bulky household waste and recyclables. This site does not accept trade waste.

If you intend to use a van to bring your waste please call 020 8509 3064 at least 24 hours prior to visiting the site. Vans are only allowed to the site between 8.30am and 4pm Monday to Saturday.

You can recycle the following items at this site:

- Bicycles
- Car batteries
- Cardboard
- Computers and printers
- Cooking oil
- Electrical items
- Engine oil
- Food/drink cans and foil
- Furniture in good condition
- Garden waste
- Gas bottles
- Glass bottles and jars
- Hardcore/rubble
- Household batteries
- Long life light bulbs
- Metal
- Mobile telephones
- Paper
- Plasterboard
- Plastic bottles
- Reusable paint
- Tetra paks
- Textiles and shoes
- Tools for reuse

TRAVEL TIMES

Oxford Circus

41 mins

42 mins

Tottenham Court Road

37 mins

47 mins

King's Cross

30 mins

35 mins

Victoria

51 mins

45 mins

Manchester

2:44 hours : mins

Paris

3:00 hours : mins

Edinburgh

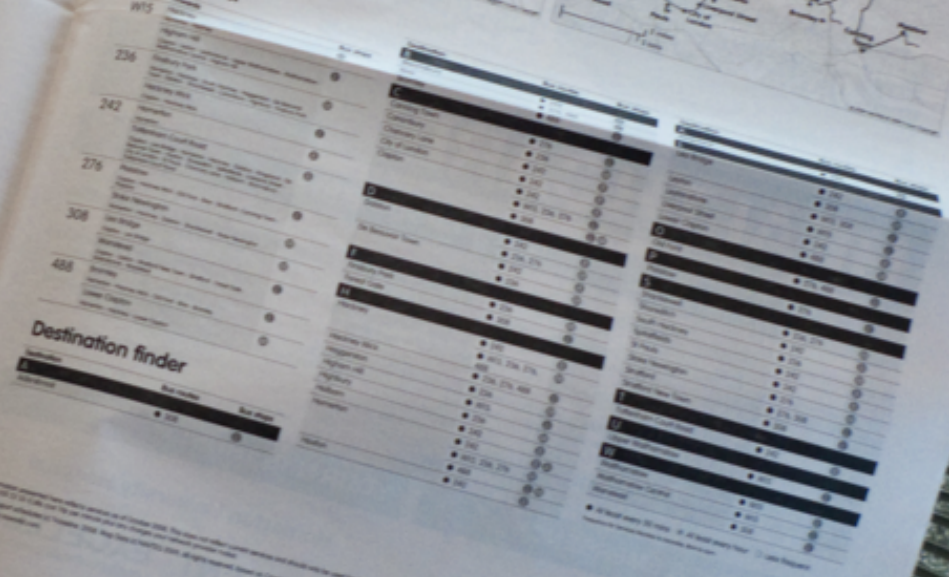
5:00 hours : mins

Bristol

3:00 hours : mins

LOCAL TRANSPORT

Route finder



Destination finder

For free cycle training in Hackney, call 020 8356 2897

ito!

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Why does this matter to you

- Supports information interoperability
- The Sum is greater than the parts
- Exploitation of
  - open data
  - open standards
- Can apply to data intranets as well as extranets
- Publishing, news, media, content services, manufacturing, retail, supply chains, finance, all classic examples where there is heterogeneous information that could be shared, linked, mapped and enriched using this approach

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Thank you

- Any further information

  nrs@ecs.soton.ac.uk or wh@ecs.soton.ac.uk

- For advice and help on applied and practical linked data projects see also

www.seme4.com