# Data, Government and the Open Web

Dr Leslie Carr,
Web Science Doctoral Training Centre & EPrints Open Access Software Centre
Web and Internet Science Research Group,
University of Southampton, UK

Introduction

- Computer Science lecturer
- Research interests: hypertext, knowledge management, open access, open data, open education, extended cognition, distributed information systems
- The Web!
- Open Source software supporting Open Access to Scientific and Scholarly Knowledge
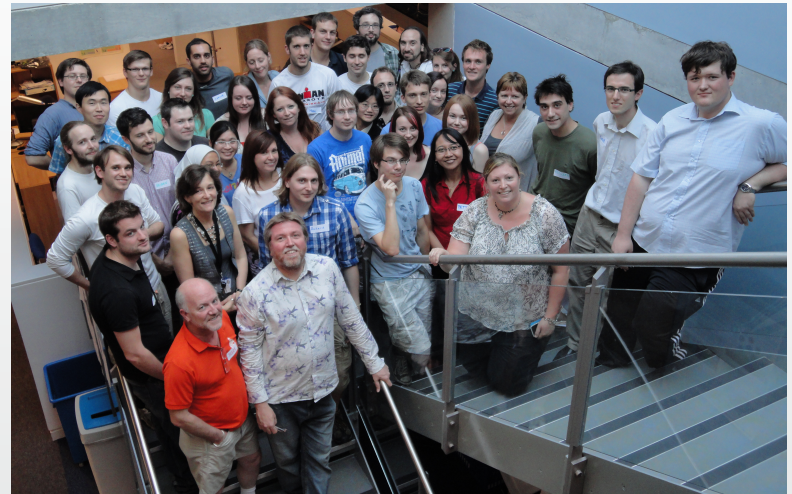 - EPrints
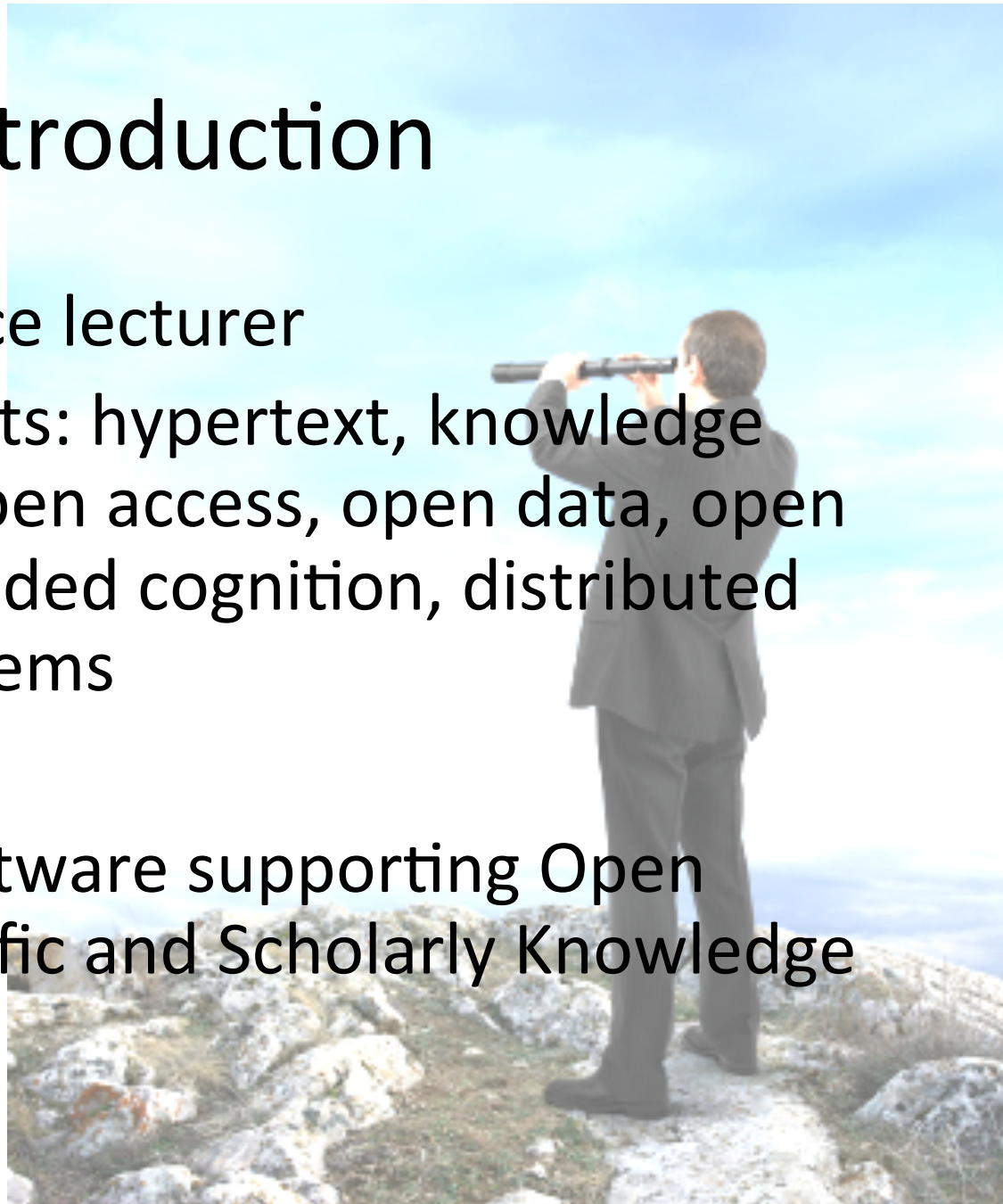
Web & Internet Science Research Group

- 20 faculty, 70 PhDs, 60 masters
- Masters courses in
  - Web Science
  - Web Technology
- Doctoral Training Centre

Why Web Science Matters

The Web has become an essential part of humanity

an essential part of the current and future economy, science and technology

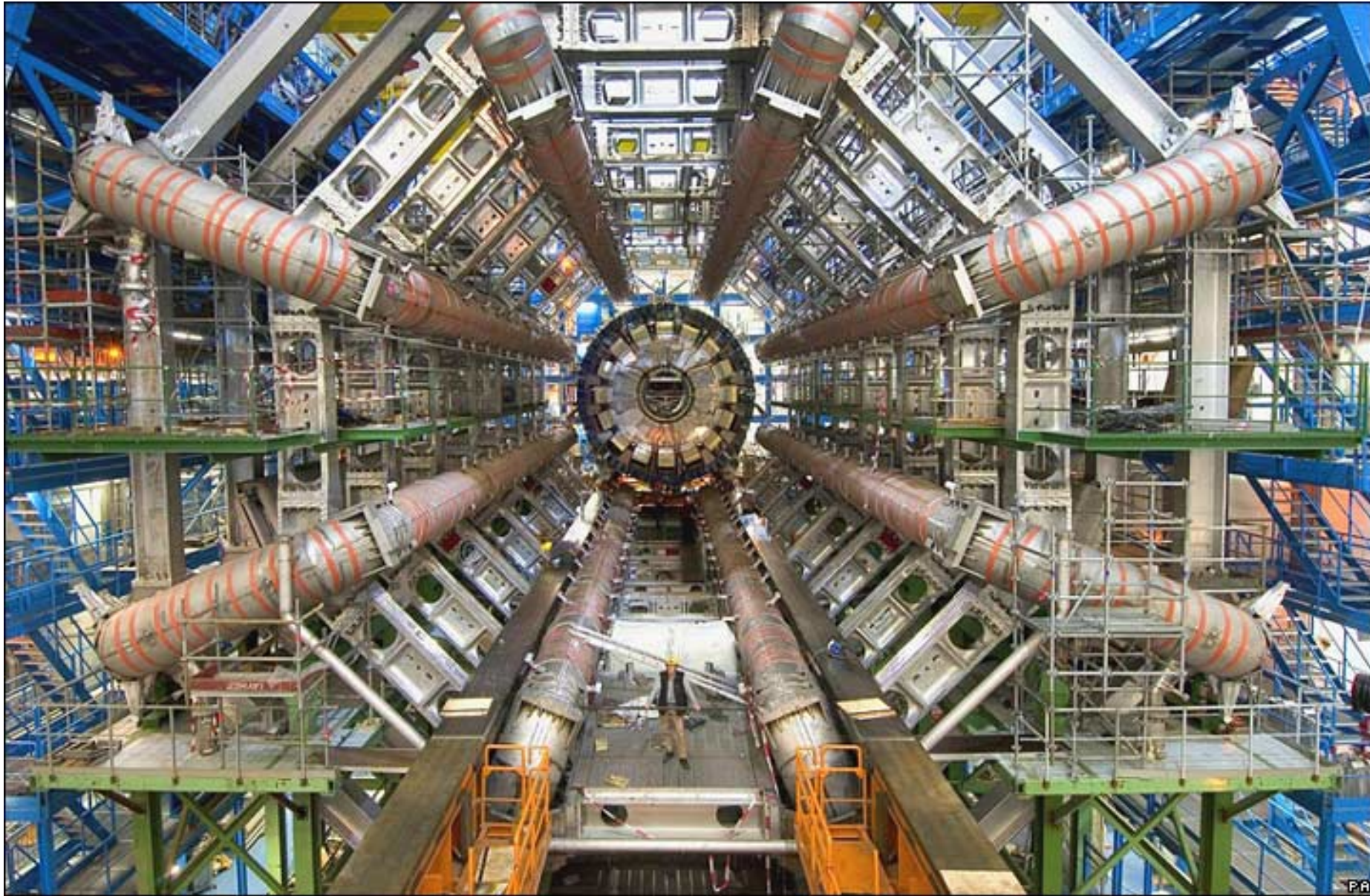- *Understanding the Web is a challenge as big as any other global cause*
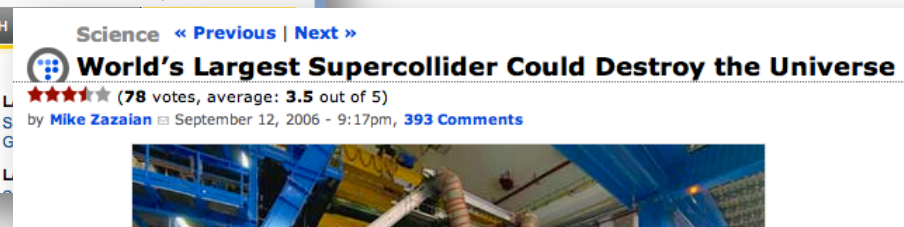
Quiz: Where's This?

Quiz: What's This?

Remember These 2006 Headline?

msn Hotmail More ▾ bing   TODAY Nightly News Meet the Press Dateline M

Home   US   World   Politics   Business   Sports   Entertainment   Health   Tech &

Space on 🦚**msnbc.com**

Good news! Black hole won't destroy Earth

Fears raised collider would create black holes that could swallow planet

NATIONAL GEOGRAPHIC   ▶ SEARCH   Site Index Subscriptions Shop Newsletters About the Society

HOME
ANIMALS
DAILY NEWS
ENVIRONMENT
GAMES
GREEN GUIDE
HISTORY
KIDS
MAPS
MUSIC
PHOTOGRAPHY

NATIONAL GEOGRAPHIC NEWS
REPORTING YOUR WORLD DAILY

Thursday, October 28, 2010

MAIN ANIMAL NEWS ANCIENT WORLD ENVIRONMENT NEWS CULTURES NEWS SPACE/TECH

Worst Case: Collider Spawns Planet-Devouring Black Hole

Anne Minard
for National Geographic News

Science « Previous | Next »
World's Largest Supercollider Could Destroy the Universe
★★★☆ (78 votes, average: 3.5 out of 5)
by Mike Zazaian ✉ September 12, 2006 - 9:17pm, 393 Comments

Birth of the Web

- From an underground nuclear bunker on the Swiss border, something did escape

- In 1989 the Web took over the academia, industry and the rest of the world

Less successful Webs from history

Sponsor	System	Scope	Real	Date	Important Properties
Finance / Press	**Reuters**	Professional, centralised	✔	1850	News & stock information (originally carrier pigeon and subsequently telegraph)
Private Institution	**Mundaneum**	Public, centralised	✔	1920	Based on indexing technology (the library card)
Military	**Memex**	Scholarly, individual, centralised	✗	1945	Aimed at Scientists and Technologists in WWII
Media	**Xanadu**	Public, decentralised	✗	1960	Focused on DRM, reuse and writing for "creatives"
Media	**CEEFAX**	Public, national, centralised	✔	1970	Broadcast, linked, not participatory
Government	**Minitel**	Public, national, centralised	✔	1980	Commercial services and information
Academy (CS & HEP)	**FTP / Archie / Anarchie**	Public, decentralised	✔	1985	Downloaded resources (papers, reports) to hard drives and printed them on LaserWriters.
Commerce	**Hypercard, HyperTIES**	Private, centralised	✔	1988	Personal applications, sometimes tied to multimedia resources on CDROMs / video disks
Academy (HEP)	**WWW**	Public, global, decentralised	✔	1990	Universal naming, linking, interoperability, participative. However no writing, no indexing.
Academy (CS)	**Microcosm**	Private, centralised	✔	1990	Sophisticated linking and openness for personal information stores
Academy (CS)	**HyperG**	Public, centralised	✔	1990	Extension of Web for with support for writing, indexing and consistency management.
Commerce	**AOL, CompuServ**	Public, centralised	✔	1990	Dialup access to email, forums, chat rooms and information resources

Society is Diverse. One Web Fits All?

Institution	Objective
Academy	Create and transmit knowledge
Commerce	Make and trade goods
Press	Investigate and report news
Media	Create and broadcast content
Police	Maintain order and public surveillance
Judiciary	Apply law and resolve disputes
Government	Control society and share resources

The development of society as a whole (nuanced and structured and refined) is inextricably related to the technology of information provision, consumption and dissemination (e.g. writing, reading, printing, education). *Different parts of society have different objectives and hence incompatible Web requirements, e.g. openness, security, transparency, privacy.*

Expansion of the Web

- The Web spread the conditions of its initial creation throughout the whole of society as it underwent an initial inflationary phase.

- The academy
  - government patronage
  - large-scale co-operation
  - sharing of intellectual property
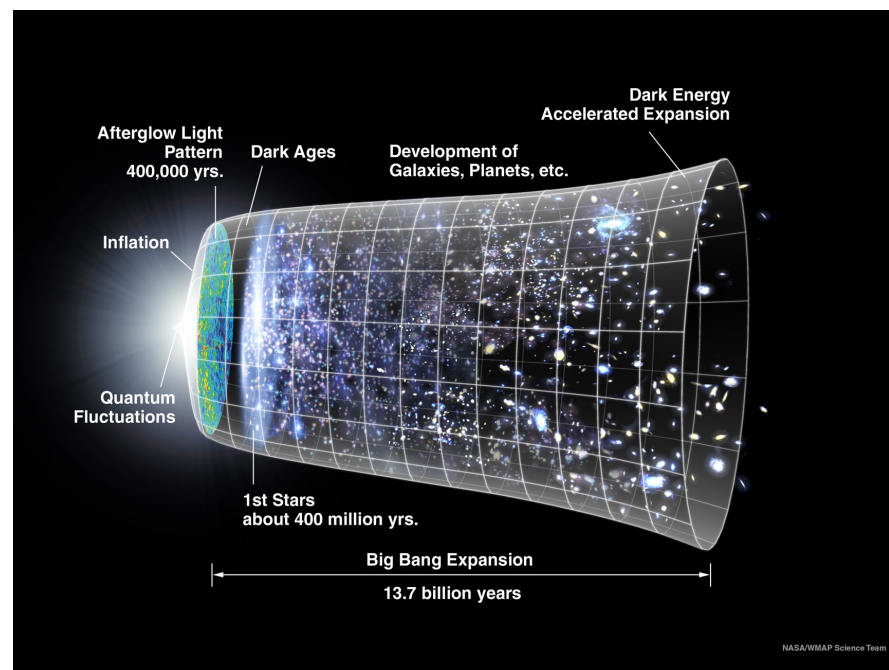
Openness Agendas

- Open Access to Research Outputs
  - Supported by all UK Research Councils
  - Heavily promoted by JISC

- Open Research Data
  - Demanded by government, private funders & UK Research Councils
  - Move from closed VRE / EScience environment

- Open Educational Resources
  - Move from closed VLEs to sharing resources

Open Research Publications
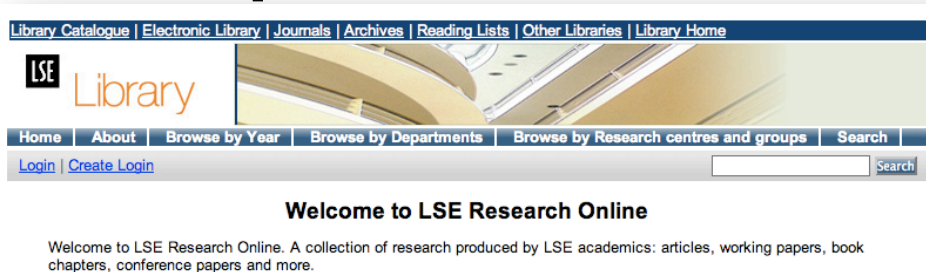
Institutional repository for the London School of Economics
http://eprints.lse.ac.uk/

One of 145 UK Institutional Research repositories
- Norway 47
- Sweden 58
- Denmark 9
- Finland 9

Open Scientific Data

A repository of chemical data JISC EBank project (2003)

Scientific, not bibliographic, metadata.

Run by the UK Crystallographic Service and latterly an international consortium.

Open Educational Resources

- JISC EdSpace Project
- IEEE LOM metadata
- Web 2, informally tagged
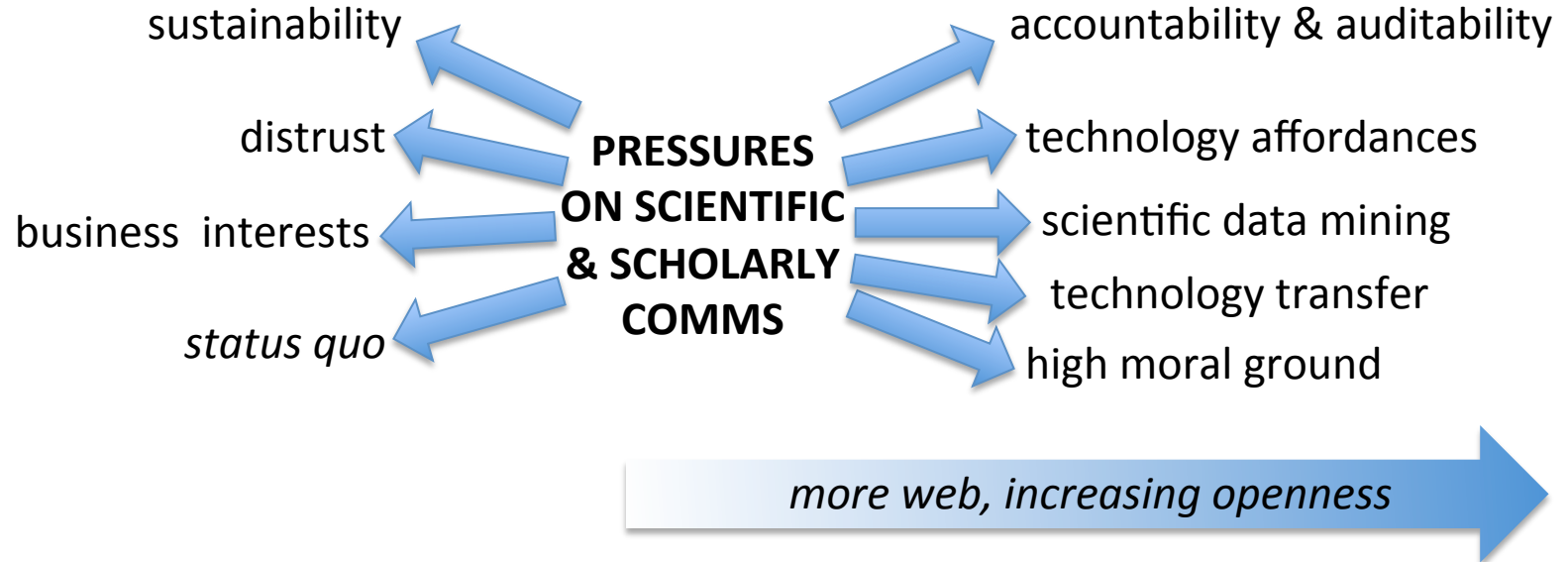- Sufficient for discoverability and usefulness

JISC Jorum (full LOM) &
JorumOpen (lightweight OA approach, DSpace)

Directions for Researchers

sustainability

distrust

business interests

status quo

PRESSURES ON SCIENTIFIC & SCHOLARLY COMMS

accountability & auditability

technology affordances

scientific data mining

technology transfer

high moral ground

more web, increasing openness

- Open Access, Open Data, Open Educational Resources
  - Web Technology joins the High Moral Ground
- *vs* the established economic model for ensuring continuity of information production
  - trading of privately held information through subscription products such as journals or magazines

Climate science must be more open, say MPs

By Roger Harrabin
Environment analyst, BBC News

MPs investigating the climate change row at the UK's University of East Anglia (UEA) have demanded greater transparency from climate scientists.

The Commons Science and Technology Committee criticised UEA authorities for failing to respond to requests for data from climate change sceptics.

31 March 2010

Roger Harrabin
Environment analyst, BBC News

Climategate

- Climate Research Unit at UEA refused to provide research data upon which its publications had been based

"It is not standard practice in climate science to publish the raw data and the computer code in academic papers," its report said.

"We consider that climate scientists should take steps to make available all the data that support their work, including raw data... and full methodological workings, including the computer codes."

"When the prices to pay are so large, the knowledge on which these kinds of decisions are taken had better be right," the report stated.

Open Government Data

- ## Same principles work on the government
  - Transparency
  - Openness
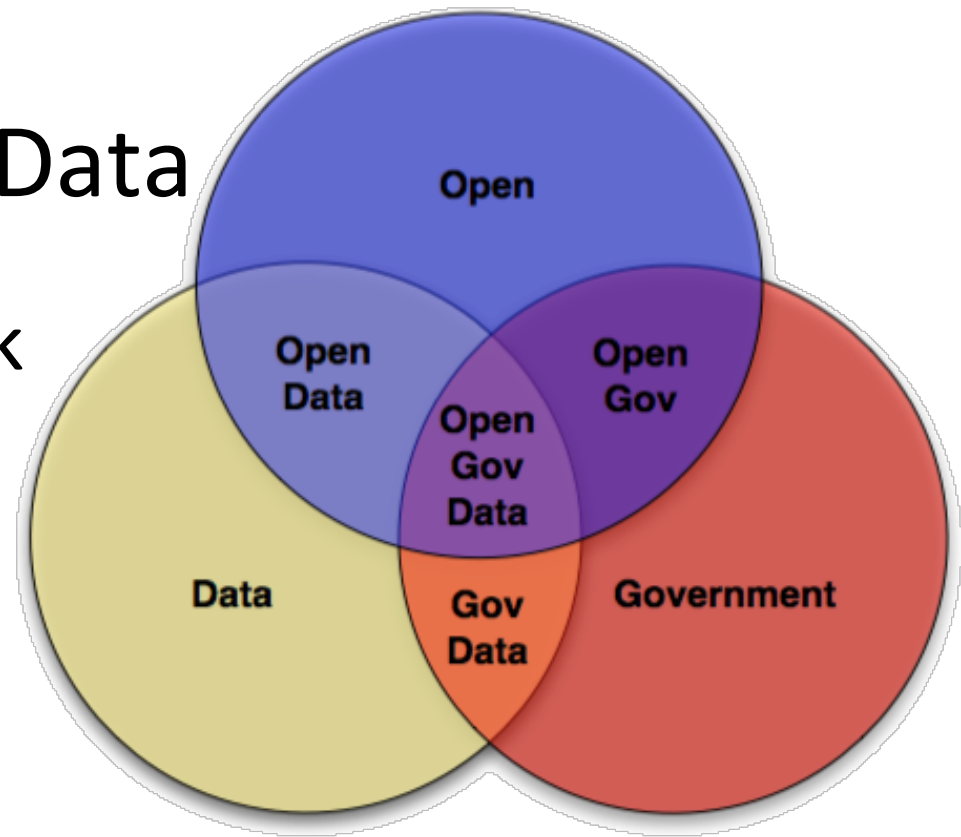 - Innovation
  - Enabling new value & new businesses
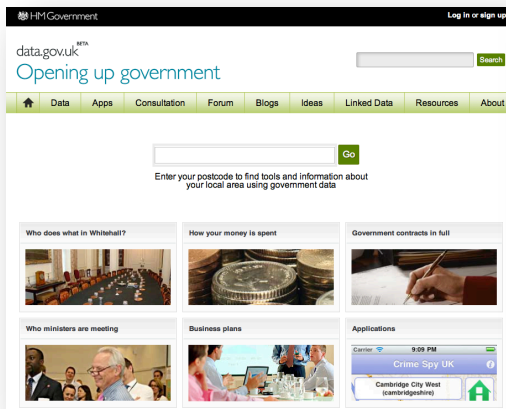
data.gov.uk (8,300 data sets)

Bespoke apps produced by independent developers for government data

Shared ideas: Michael built a 'planning proposals' web app. Some turned it into an RSS feed. Someone else turned that into a Twitter feed. Someone else used scraperwiki…

Data.gov.uk - What's Available

-Over 8,300 datasets now
-Frequent deposit activity of medium sized dataset deposits
-Signs of a healthy data catalogue

Dataset Deposits made in Data.gov.uk
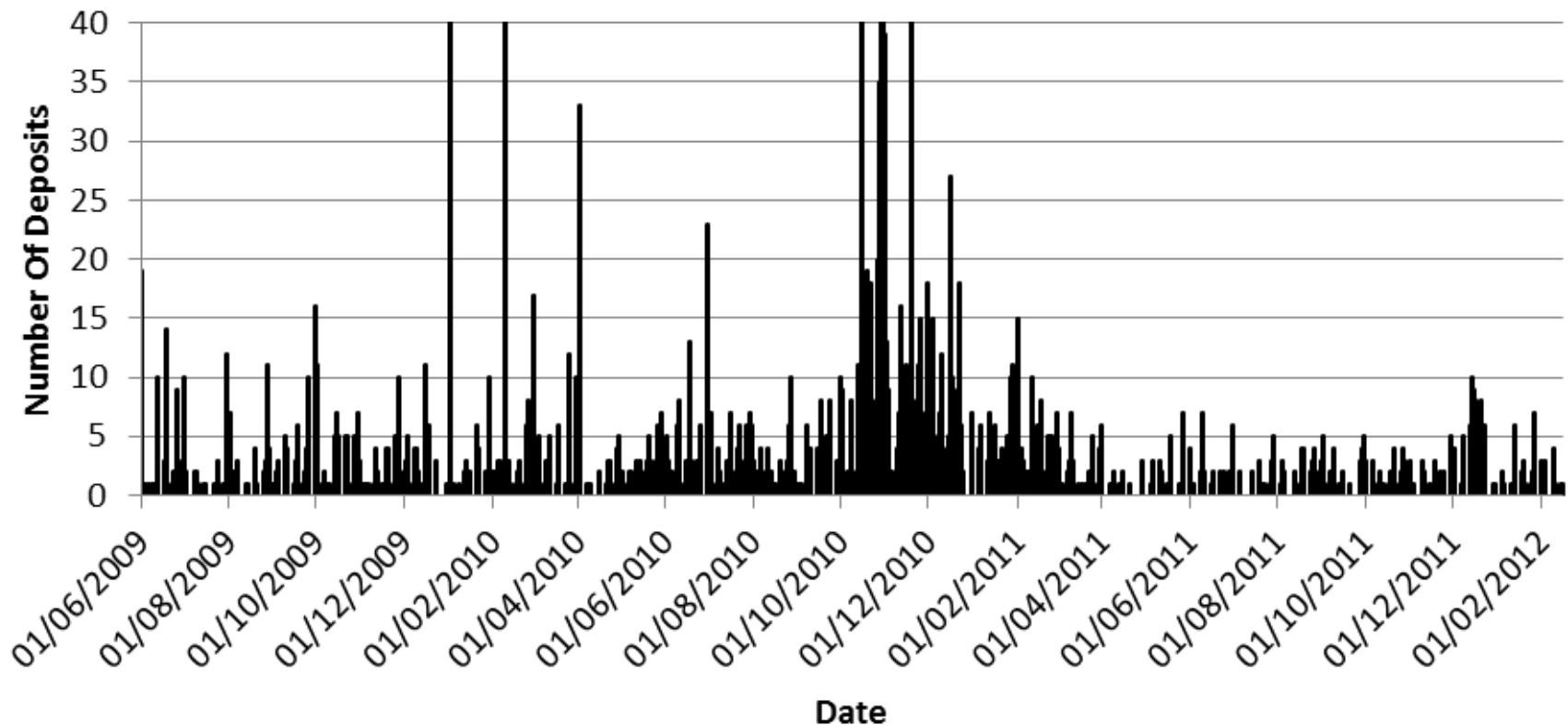
Data.gov.uk - Who's Publishing It

- 406 departments supplying data
- These include national Government departments, local government, and local council

Departments Depositing >1% of Total Datasets in data.gov.uk

- Department of Health [11410] — 26%
- Department for Communities and Local Government [11401] — 20%
- Office for National Statistics [11606] — 10%
- Department for Environment Food and Rural Affairs [11403] — 7%
- Department for Children Schools and Families [11398] — 7%
- Home Office [11402] — 5%
- Ministry of Defence [11404] — 4%
- Department for Business Innovation and Skills [11399] — 4%
- Department for Work and Pensions [11409] — 4%
- Ministry of Justice [11400] — 3%
- NHS Information Centre for Health and Social Care [16121] — 2%
- Department for Transport [11406] — 2%
- Department for Culture Media and Sport [11419] — 2%
- Department for Education [11424] — 2%
- Warwickshire County Council [11428] — 2%

Data.gov.uk - Who's Using It

- 197 Applications available on Data.gov.uk
- Ranging over multiple topics of interest
    - Transparency Apps
    - Transport Apps
    - Public Health Apps

Number of Open Data Applications Published on data.gov.uk

Series1	2009	2010	2011	2012
	6	102	75	9
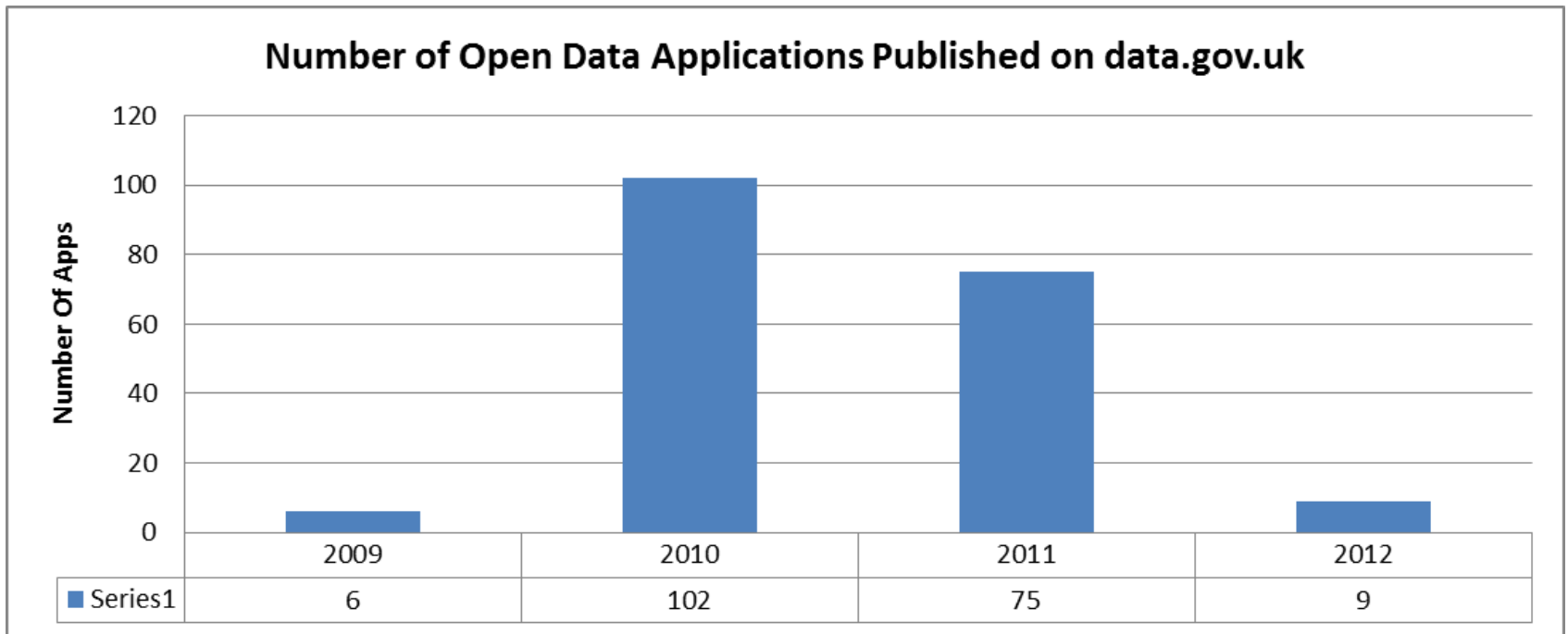
Institutional Open Data

@base <http://id.southampton.ac.uk> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

<http://data.southampton.ac.uk/dumps/places/2012-01-09/buildings-ene
    <http://purl.org/void/provenance/ns/processType> </ns/DownloadV:
    <http://purl.org/void/provenance/ns/resultingDataset> <http://da
    <http://purl.org/void/provenance/ns/sourceDataset> <https://docs
hl=en_US&hl=en_US&key=0AqodCQwjuWZXdHdoZEhCdXRtR2VEMFQ1NWNnbWpKZmc&
    a <http://purl.org/void/provenance/ns/ProvenanceEvent> ;
    <http://www.w3.org/2006/time#hasBeginning> "2012-01-09T20:55:41:
    <http://www.w3.org/2006/time#hasEnd> "2012-01-09T20:55:42Z"^^<ht

<http://data.southampton.ac.uk/dumps/places/2012-01-09/buildings-est
    <http://purl.org/void/provenance/ns/processType> </ns/DownloadV:
    <http://purl.org/void/provenance/ns/resultingDataset> <http://da
    <http://purl.org/void/provenance/ns/sourceDataset> <https://docs
hl=en_US&hl=en_US&key=0AqodCQwjuWZXdGdySXpZaS12UH1HSTF0Ni10akNMbWc&
    a <http://purl.org/void/provenance/ns/ProvenanceEvent> ;
    <http://www.w3.org/2006/time#hasBeginning> "2012-01-09T20:55:39:
    <http://www.w3.org/2006/time#hasEnd> "2012-01-09T20:55:40Z"^^<ht

<http://data.southampton.ac.uk/dumps/places/2012-01-09/buildings-ext
    <http://purl.org/void/provenance/ns/processType> </ns/DownloadV:
    <http://purl.org/void/provenance/ns/resultingDataset> <http://da
    <http://purl.org/void/provenance/ns/sourceDataset> <https://spre
ylnmhsrdCladGZrVlNlQnZ22RXNMcnpMalRfNlE&single=true&gid=2&output=csv:
    a <http://purl.org/void/provenance/ns/ProvenanceEvent> ;
    <http://www.w3.org/2006/time#hasBeginning> "2012-01-09T20:55:41:
    <http://www.w3.org/2006/time#hasEnd> "2012-01-09T20:55:41Z"^^<ht

<http://data.southampton.ac.uk/dumps/places/2012-01-09/buildings-po:
    <http://purl.org/void/provenance/ns/processType> </ns/DownloadV:
    <http://purl.org/void/provenance/ns/resultingDataset> <http://da
    <http://purl.org/void/provenance/ns/sourceDataset> <https://spre
ylnmhsrdCladGZrVlNlQnZ22RXNMcnpMalRfNlE&output=csv> ;
    a <http://purl.org/void/provenance/ns/ProvenanceEvent> ;
    <http://www.w3.org/2006/time#hasBeginning> "2012-01-09T20:55:40:
    <http://www.w3.org/2006/time#hasEnd> "2012-01-09T20:55:40Z"^^<ht

<http://data.southampton.ac.uk/dumps/places/2012-01-09/buildings-sha
    <http://purl.org/void/provenance/ns/processType> </ns/DownloadV:
    <http://purl.org/void/provenance/ns/resultingDataset> <http://da
    <http://purl.org/void/provenance/ns/sourceDataset> <https://spre
ylnmhsrdCladGZrVlNlQnZ22RXNMcnpMalRfNlE&single=true&gid=1&output=csv:
    a <http://purl.org/void/provenance/ns/ProvenanceEvent> ;
    <http://www.w3.org/2006/time#hasBeginning> "2012-01-09T20:55:40:
    <http://www.w3.org/2006/time#hasEnd> "2012-01-09T20:55:41Z"^^<ht

<http://data.southampton.ac.uk/dumps/places/2012-01-09/places.rdf#p:
    <http://purl.org/void/provenance/ns/processType> </ns/ConvertTu:
    <http://purl.org/void/provenance/ns/resultingDataset> <http://da
    <http://purl.org/void/provenance/ns/sourceDataset> <http://data
    a <http://purl.org/void/provenance/ns/ProvenanceEvent> .

Institutional Apps

Application created by 4 students using the following datasets:

- Buildings and Places
- Catering
- Local Amenities
- Southampton bus-routes
- Southampton Bus-stops
- Teaching Room Features
- Ordnance Survey Data

How Does Open Data Work?

- Make catalogue of all the things **you** know

- Give each thing a permanent web name (URL)

★ Publish information about that thing on the Web with an open license

★★ Use a machine readable format

★★★ Use a standard data format

★★★★ Use an open standard data format

★★★★★ Mention other things using their web URLs

"Open" Platforms

Internet

"Open" Platforms

Internet

Web

"Open" Platforms

Internet

Web

| Wiki | Facebook | Twitter | YouTube | Google |

"Open" Platforms

Internet

Web

Wiki	Facebook	Twitter	YouTube	Google

Wikipedia	Social Marketing	Media Channel	**X**

Why Closed Knowledge Systems are Bad

Evaluation

Did Germany make the right choice in relocating their post-unification capital to Bonn?

Synthesis

Create a set of guidelines for choosing alternative national capitals in a post-climate change Europe.

Analysis

What were the overriding factors in the choice of sites for the European parliament and commission?

Application

Choose the most appropriate capital for the Isle of Wight
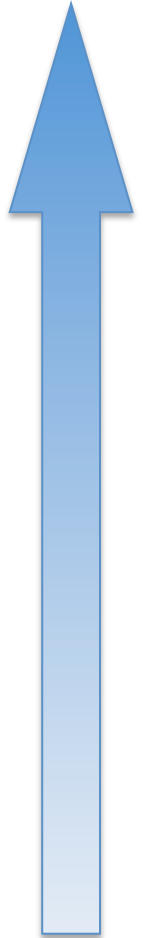
Comprehension

Why is Paris the capital of France?

Knowledge

What is the capital of France?

Increasing educational capabilities

Bloom's Educational Taxonomy demonstrates why closed knowledge systems are bad

Open Source, Systems, Data + Us = Open Society

- The Web isn't a thing but an activity
  - the creation of a network of information by a network of individuals.

- The Web wasn't invented by Tim Berners-Lee,
  - it is still being invented by all of us as we gradually adapt our tools and change our practices

The Web is an extraordinary change in the ability of humankind to build value and to be valuable
The product of people with open license, open standards, open systems

Co-constitution