data.gov.uk - the linked data revolution

Nigel Shadbolt
UK Government Information Advisor

SAP
24 March 2010
data.gov is taking hold
Unpacking innovation
Working with UK Public Sector information and data

Advised by Sir Tim Berners-Lee and Professor Nigel Shadbolt and others, government is opening up data for reuse. This site seeks to give a way into the wealth of government data and is under constant development. We want to work with you to make it better.

We're very aware that there are more people like you outside of government who have the skills and abilities to make wonderful things out of public data. These are our first steps in building a collaborative relationship with you.

Latest news:
- Read about our latest site changes
- Information about the Local Data Panel
- Publications in the news

Search Data
Enter keyword(s)  Search
- e.g. education, NHS, crime, transport, environment
- Powered by: CKAN

Browse for Data
List all datasets
By Public Body

Most Recent Apps
- SaferMK
  This is a crime reduction application that has two key functions: 1. simple, broken down interface.
- EnAktng UK Public Sector Information Data Mashup
  Map application that using OpenSpace
- OpenGeoscience
  Discover if you live on rocks from an extinct volcano, in the middle of an ancient river or forest

Digital Engagement
Twitter stream
@dataandthings: UKatsbybritain Announced in House today establishment of Broadband Delivery UK, chief executive Adrian Kemelver Follow us on Twitter
Tell us your idea
MPDP: Making Public Data Public

- Berners-Lee and Shadbolt - June Appointed HMG Information Advisors
- Context - Pol Taskforce
- Previous work with OPSI
- Strategic Drivers
  - Transparency and accountability
  - Economic and Social Value
  - Public Service Improvement
  - New Industries New Jobs
Key Elements 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 information regulatory framework
Progress 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
Accomplishments

• Embodied Public Data Principles in Government
• Beta service launched data.gov.uk - Jan 2010
  Currently >3000 datasets including linked data
  2500 developer community
  Open source and open standards
  Self describing Catalogue
• Key data sets under consultation for release
• Applications emerging
• New Crown Commons data licence
• Local Data Panel launched
• Cross party interest and endorsement
Our public data principles state that:

1. Public data will be published in reusable, machine readable form
2. Public data will be available and easy to find through a single easy to use online access point data.gov.uk
3. Public data will be published using open standards and following the recommendations of the World Wide Web Consortium
4. Any ‘raw’ dataset will be represented in linked data form
5. More public data will be released under an open licence which enables free reuse, including commercial reuse
6. Data underlying the Government’s own websites will be published in reusable form for others to use
7. Personal, classified, commercially sensitive and third-party data will continue to be protected.
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.
data.gov.uk
Wouldn’t it be good if...?

Ideas – the seeds of innovation

Got a great idea you’d like to see realised? Wouldn’t it be good if, for example, you could see a map that showed all the local schools, alongside their performance ratings for instant comparison? You might not be able to create such a map but you know it would be useful. Submit your ideas here and the chances are high that someone with the know-how, who shares your interest, will pick up your suggestion.

Submit your idea

Most Recent Ideas

- County Court Judgement Data
  County Court Judgement data is a matter of public record, anyone willing to pay per search is welcome to the data.
  By Phil
  21/01/2010 - 11:32

- Fuel prices vs public transport provision by area
  It would be useful to see how the cost of fuel affects people in rural or remote areas where public transport may not be as prevalent or useful as in urban areas.
  Showing...
  By Laura
  21/01/2010 - 11:24
data.gov.uk
ASBOrometer is a mobile application that measures levels of anti-social behaviour at your current location (within England and Wales) and gives you access to key local ASB statistics.

ASBOrometer is available for iPhone and Android phones. Get it FREE from the iTunes App Store or Android Market now!

This app was created by Jeff Giffelt and made possible by the data.gov.uk initiative, which is opening up UK government data for public reuse.
UK Dentists

Find your nearest NHS dentist quickly and easily!

The National Health Service has over 7,500 dental surgeries across England but finding one can be a difficult and time consuming task. This app is based on data provided by the UK's Health and Social Care Information Centre (HSCIC)* and covers 99.4% of all registered NHS dental surgeries in England.

Search for an NHS dentist around your current location, or look for one in another area (where you may be moving to, or for a friend) via a simple place name or postcode search. Distances can be displayed in metric or imperial units (configured in the application preferences).

UK Dentists uses the built-in GPS on the iPhone to get your current location, but can also locate an iPod Touch in built-up WiFi areas. If your location cannot be determined automatically then you can always enter your postcode to get started. Please note that an internet connection is required to get location and map data.

* Data reproduced under the terms of the Office of Public Sector Information (OPS) Click-Use Licence.

Application Screenshots

Click on any of the images to see a larger version

Buy on iTunes
Massive Outgoings
Cost for average British taxpayer per day

GOVERNMENTS

£2.93
£1.33
£3.26
£5.92
£0.13

ESSENTIALS
NHS
£9.05
Schools
£2.56
Defence
£3.99
TRAVEL
Roads
£3.10
BANKING CHARGES
LRT
£0.14
Recapitalisations
£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
**NEIGHBOURHOOD**

![Graph showing crime rates over time]

Summary
---
Hackney Police and the Council work very closely.

Police Confidence: 77% up 39%
Overall Crime: Down 37% - 17%
Gun Crime: Down 14% - 13%
Knife Crime: Down 220% - 10,000
Robbery: Down 20%
Theft: Down 37%

In 2008-09, gun crime was reduced by over 37% and knife crime by 14%. Youth crime also fell by 10%. Satisfaction with the Police and the Fire Brigade are six per cent above the London average.

---

**GOLDEN SCHOOL**

- **Alotments**
  - Tower Hamlet: 282
  - Waltham Forest: 1,382

- **Waiting List Time**
  - Tower Hamlet: 415
  - Waltham Forest: 113
  - Newham: 1,715
  - Hackney: 195

**GREEN SPACE**

- 23% of the area is Green Space

**RECYCLING**

- **% Satisfaction with Parks**
  - Hackney: 90%
  - London average: 80%

- Recycling and Waste Centre:
  - Open Monday to Saturday, 9am to 5pm.
  - Accepts paper, plastics, glass, and metal.

You can recycle the following items:
- Textiles and shoes
- Tools for reuse
- CDW (Construction and demolition waste)
New data?

• The objective of the Public Data work is to make data available for re-use.

• The data on data.gov.uk includes:
  – data not previously available at all
  – data previously only available in non-reusable form (eg PDF)
  – data previously only available with licensing restrictions on its reuse (there
    is a new, open licence for data on data.gov.uk)
  – data previously only available to a limited or specialist set of users
  – data previously not easy to find
What else needs to happen in HMG

• each body needs a clear, published commitment on releasing data, and there needs to be consistency between similar bodies

• agencies need to release transaction data - like the bicycle accident data - and not just statistical aggregates

• this needs to extend not only through central Government but also into local government and the other public services

• agencies need always to publish for re-use the underlying data for web services that they supply like Directgov, NHS Choices and the Camden "where's my nearest" service.

• bodies should proactively release data, but there should also be the ability to challenge bodies to release more.
Nuturing the community

• Government also needs to nurture and support activities which would add new value to data.

• Support community activism and "hyperlocality" in using this information.

• Help integration with community networks - e.g. civic society services.

• Assist the use of linked data technologies by investing in the development of tools and software and making these "open source".

• Giving support, encouragement and targeted funding to people developing innovative software and services which use public data.
Challenges we faced

• Cultural & Organisational
  – Data hugging
  – We have released it - sort of…
  – License impediments

• Worries about
  – Confidentiality
  – How it will be Interpreted
  – QoI
  – Lost value
  – Disrupted workflow
  – New IT procurement

• Importance of local authority and other public bodies

• Capability
  – What assets have we got (no duty to hold register of public non-personal datasets)
  – Who can do this work
  – How to share best practice
  – Minting URIs

• Policy
  – Extending the principle to all public sector bodies
  – Consistency in what is published, how it is licensed and its format
  – Range of exceptions do need to be appreciated
  – Confidentiality, accuracy, liability and reputation
The US Linked Data Effort

Data Processors
(Web Services & Analyzers)

- XSLT Service
- Diff Service
- Link Annotator
- RSS Generator

SPARQL End Point
SPARQL Web Service

Linked Data
(RDF)

GOV data
(RDF)

CSV
XSL
...

Convert Data
View & Use Data
Link & Enrich Data

Sem Wiki
Tabulator
Google Viz
MIT Exhibit
RSS 1.0
tagCloud

Data
Processors

Li Ding, Dominic DiFranzo, Sarah Magidson, and Jim Hendler · Tetherless World Constellation · Rensselaer Polytechnic Institute · Aug 7 2009 · http://data-gov.tw.rpi.edu/
This application presents a mashup of foreign aid data (represented in US Dollars) from the United States Agency for International Development (USAID) and UK Department for International Development (DFID) for the 2007 US Fiscal Year.

This demo uses data from Dataset 311 on crime in the United States, census.gov and UK recorded crime. It shows available information on crimes in cities grouped by US countries and UK Territorial Police area.
Household energy usage statistics.  

Energy usage (by energy type) of US government agencies.  
Comparing Smoking Quit Rates to Unemployment, Tax, and Violent Crimes

This demo aims at showing how existing tobacco survey data can be linked to other open government data in unexpected way.

http://data-gov.tw.rpi.edu/demo/linked/2010/02/04/popscigrid/nih.html
....It is equally hazardous trying to envision how freer data will redraw the boundaries between different communities or recast their relationship with power. But it is reasonable to speculate that the uncovering and unlocking of so much information will drive improvements in public policy. It will level the territory on which voters meet politicians, and could prove a powerful brake on campaigning hyperbole in the coming election. Without the printed word there would have been no informed electorate, no demand for accountability from our leaders – and indeed no democracy at all. Open data will surely revive it, and in time could transform it too.
Web of Linked Data

- The Web of Linked Data is next
- Can change
  - Government
  - Business
  - Cities
  - Universities
  - :
- Exploitation of
  - open data
  - open standards
- Can apply to data intranets as well as extranets
So the question is...

If you could get the data under Freedom of Information request why shouldn’t it be published as an RSS stream?

The question is not why but why not?
Thank you

• Any further information or follow up on the research questions please email

nrs@ecs.soton.ac.uk