

SWIRRL IT RESPONSE (VIA E-MAIL)

There are now thousands of datasets on data.gov.uk, most of them CSV files/Excel/PDF etc. It's a good first step, but we need to help people find relevant information and to use it effectively.

* Linked Data makes the data more powerful, more useful and more findable by interconnecting different datasets. It might not be appropriate for all datasets on data.gov.uk, but in many cases it will greatly enrich them.

* the government has a role to play in encouraging and enabling data providers to move from '1-3 star' data to 4 and 5 star data. Para 8.9 proposes this as an option. The government should take steps to help to make it happen.

* Some very useful guidance documents, URI sets and ontologies have been created already. But more can be done to build out the 'data infrastructure' and establish key datasets for interconnections. The administrative geography, date reference are a great start, and the recent Companies House linked data is an excellent addition - but there are many more such reference data sets which could be made much more useful using the Linked Data approach.

* Some suggestion for dataset which would have particular value if published as Linked Data:

- 'naming of parts' URIs for anything important - all government physical assets (buildings etc)
- Ordnance Survey, statistics.data.gov.uk and location.data.gov.uk are an excellent start on geographic data, but also consider:
 - other ONS regions (Census Output Areas etc),
 - the road network model - URIs for nodes and edges of our road network
 - better geographic look-up services (like mapit.mysociety.org but better and Linked Data-based)
- demographics
- more on public transport
- the National Land and Property Gazetteer
- land use and ownership
- economic data (recent and fine-grained)

* Ordnance Survey 'derived data' licence restrictions have a very substantial negative impact on the ability of the UK public sector to open up more useful data. The work of local authorities in particular and their ability to be transparent could be greatly simplified if this licensing question was properly addressed. Much time is wasted trying to understand if their OS licences allow them to do something or not.

* There is the potential to draw more effectively on the expertise of Linked Data practitioners in the UK and worldwide. We have contributed to a separate response to the consultation from a group of Linked Data practitioners where the specific Linked Data issues are presented in detail.

* Data typically gets from source to end user via expert intermediaries - application or visualisation developers, reseearchers, journalists etc. The government could help to demonstrate best practices to these intermediaries to help them make effective use of the available data, and to show data providers how to make life easy for the intermediaries.

* The government should consider putting in place steps to ensure the scalability and longevity of the data infrastructure. This could involve reverse proxies for distributing the different parts of the data.gov.uk domain to various data stores and/or a PURL service of some sort.

* Organisations putting open data publishing in place should bear in mind the responsibility to keep data accurate and up to date, so that users of the data can put their trust in it.

* Productivity benefits of open data. The consultation document puts a lot of emphasis on use of open data for cost monitoring/comparison and benchmarking of outcomes. That is indeed useful and a true benefit, but another very important economic benefit of open data is reducing the costs of data search and data integration when different parts of the government have to work together. Open data (and in particular linked data) can have a significant impact on this, leading to improved efficiency in delivery of public services.

* The National Archives should consider the issues around collecting and archiving the most important government open datasets.

* the scale of open data is going to keep growing. Government should put money into research into the technologies for improved access to large scale distributed data - perhaps through Technology Strategy Board?

* there is a need for better 'data education', in schools and universities and 'continuing professional development' for public sector staff. Understand how to access, critically assess and use available datasets. As more data becomes available, we need people to have the skills to make use of it, and to encourage use of evidence for mature discussion. Better education on use of data throughout the population (a long term goal of course) would help counter people making false arguments by 'selective quoting'.

* Distinction between personal data and open data: giving individuals better access to their personal data held by government (as referred to in 4.2 and 5.1) is laudable but does not relate to *open* data, so should be addressed elsewhere. Including personal data in discussions of open data may confuse the issue.