

RICHARD MASON REPOSE (VIA E-MAIL)

I am responding as an individual developer at a County Council working mainly on the public website. I am an admin for our data.gov.uk account and would be likely to be involved in publishing data.

1.2 When a decision is being taken about whether to open a dataset, what tests should be applied?

It should be based on the same tests as FOI, so that there is only one set of requirements to refer to. Update the requirements for all of FOI to include the requirements of datasets.

8.1 How would we establish a stronger presumption in favour of publication that that which currently exists?

A new requirement that public bodies publish data wouldn't help enforce a right to data. It's too vague – reluctant authorities could simply publish one or two datasets and declare the requirement met.

8.2 Is providing an independent body... the most effective option...?

Establishing a right of challenge with power to enforce format, quality and regularity of publication would be more effective.

8.4 What might the resource implications of an enhanced right to data be? How do we ensure any additional burden is appropriate?

A requirement to publish any datasets on demand by a requester willing to pay could certainly prevent public bodies delivering their core functions. Releasing data may well involve a lot of work, eg developing/customising software. If the FOI response of 20 days is enforced this could easily force managers to pull developers off of scheduled projects at short notice in order to meet the FOI deadline. Teams are often working at full capacity and this could easily result in other deadlines being missed, many of which would incur their own costs.

The FOI deadline for a response should remain, but that response could be a plan for how and when the data will be provided. Actually providing the data should have a longer deadline, perhaps 6 months. Although that's a significant delay it's a realistic timescale, and data would be published on a regular schedule thereafter.

The consultation also talks about authoring data in tools which are 'online by default'. However this ambition may be limited by the requirement for data to be held within UK borders. Many cloud-based services cannot guarantee this.

8.5 How will we ensure that open data standards are embedded in new ICT contracts?

Much like the development of the Open Government Licence, central government should develop a standard clause designed to build open data into public sector ICT contracts, and then require that that clause is included in all new ICT contracts.

8.1 What is the best way to achieve compliance on high and common standards to allow usability and interoperability?

For datasets which must be published by multiple local authorities, it would help if similar datasets could be tied together so it's easy to see what formats others have used. A good way to do this would be to maintain a list of local authority datasets, each with its own id, on a similar model to the Service List maintained at <http://doc.esd.org.uk/ServiceList>. This could be called the Dataset List.

8.2 Is there a roles for government to establish consistent standards for collecting user experience across public services?

Yes, but as enabler rather than enforcer.

Government already sets some standards for data publication, eg the Egovernment Interoperability Framework, Egovernment Metadata Standard, the Government Data Standards Catalogue and BS7666 for addresses. These were set up by the now defunct govtalk.gov.uk. data.gov.uk should take ownership of these and be clear whether they are still current recommendations or deprecated.

8.1 How should public services make use of data inventories? What is the optimal way to develop and operate this?

Paragraph 8.10 talks about merging data lists into a single data inventory. The new Code of Recommended Practice for Local Authorities on Data Transparency talks about publishing that list on data.gov.uk. This would be a missed opportunity: data.gov.uk shouldn't reference a list for each public body, *data.gov.uk should be the list*.

Data.gov.uk should be a list of all datasets held by government, whether open or not. Data published is a subset. Some changes would be required to enable this:

- Published data should be presented before unpublished data
- data.gov.uk should be the "unlocking service" for the rest. Unpublished datasets should have a "request this data" button.
- If publication has already been refused under FOI rules the reason should be stated in the dataset metadata.
- Support an "other" option for licences, which might take a URL or free text.
- Publishers should be required to rate each dataset from 0-5 using the Five Star Rating for Open Data. Publishing lots of low scores would focus their attention.
- Public statistics on data.gov.uk should compare the performance of public bodies on their star ratings. Use transparency to encourage transparency.
- Datasets should be matched against the Dataset List (see response to earlier question). Then similar bodies could be compared eg are all County Councils publishing dataset #123 and at what star rating?

Data.gov.uk should work with SOCITM to include star ratings in their annual Better Connected report for local authorities.

8.5 Should the data government releases always be of high quality?

No. If there are concerns about quality they should be stated, with specific detail, in the metadata for the dataset. Use transparency to enable transparency.

8.1 How should government approach the release of existing data for policy and research purposes: where should it be held?

It doesn't matter, so long as it's registered in the data.gov.uk inventory.

8.3 Which is more important: broader sets of data or detailed data?

Similar to open source software, allow users of data to log "issues", and to vote on which issues are most important to them. Issues can include publishing a new dataset, fixing errors, or expanding detail. This mechanism allows priorities to be led by demand.

Rick Mason