

Open Data Consultation response from the Evidence and Analysis Unit at Department for Culture, Media and Sport (DCMS)

Intro

Thank you for the opportunity to respond to this consultation which we take great interest in. DCMS has been fully engaged with the government transparency agenda (as documented in our Business Plan), and have cooperated in central exercises such as HMT's data collection review, and DCLG's single data list. As a matter of principle we release a number of statistical series and data (providing it is not commercially sensitive or breaches privacy confidentiality laws), and for all official and national statistics these are in line with the Code of Practice for Official Statistics.

General Considerations

1. As a small department we do not hold a large number of data collections, but we do sponsor a number of Arm's length Bodies (ALBs) which collect and own much of the data DCMS uses. We would be interested to know to what extent the Open Data proposals would cover these ALBs. Also data is owned and mitigated within many different parts of an organisation (it is not just government statisticians), all of whom need to be aware of the requirements of making the data freely accessible. This agenda needs to be communicated to all staff from the centre.
2. There must be considerations into the extent of these proposals, and how these would not breach data privacy laws, and how to balance any economic growth initiatives through access to data against unfair competitive advantages and potentially ending working relationships with business data suppliers who do not want their details published.
3. There is a danger of opening up large datasets without any commentary and explanation, as results can be misinterpreted or misused. Also there are complexities with some datasets (especially those which are not from administrative sources, i.e. surveys) which require the user to have specialist skills or experience of using.
4. DCMS feel a key aspect to 'open data', is with data linkage between government departments. The benefits all departments could gain by linking our data sources together could be huge and could enrich analysis being conducted, boost sampling frames, and sample sizes and provide a more holistic understanding of society and not look at policy issues in departmental silos. There could also be benefits to individuals and businesses, with the potential to create social and economic growth. We therefore would welcome the open data consultation to provide easier access for departments to do this, and reduce many barriers such as legal issues.

5. There are resource and cost implications on having a 'right to data'. There must be limits, like those applied to FOI requests. In current times, we feel this could place excessive burden on data producers, and instead public bodies should be encouraged to release data, to as granular level, as is deemed sensible and appropriate. This also places an onus on data owners regularly releasing and updating the data, even if the original request for data was a one-off specific release.
6. A lot is already being done by departments, and there are a number of avenues users can use to access our data, from departmental websites, to the publication hub and data.gov to the UK Data archive. DCMS aim to publish more detailed breakdowns of data in a re-usable format (CSV) where disclosure control allows. For our flagship survey, Taking Part, we have also developed user friendly interactive tool, NetQuest, which allows data to be analysed quickly and easily without needing any knowledge of specialist software. Actions following this consultation should not prescribe a single solution for disseminating data and official statistics. Making data available in different ways and in different places helps maximise their use. Also using new technology and social media can create difficulties for departmental IT, Press and Comms units. For example at DCMS, Comms control all releases via social media channels, which would potentially cause problems with the code of practice for official statistics.

Consultation questions

1. Do the definitions of the key terms go far enough or too far?

It does not mention survey data, nor qualitative research data collections or some experimental statistics. Survey data without any identifiable attributes to the respondent can be released so could be included in your definition.

2. Where a decision is being taken about whether to make a dataset open, what tests should be applied?

Is the dataset already open to the public in some form, and if so how? Are there any personal details or market sensitive information which needs to be restricted? Can datasets be released after protection control procedures have been applied to suppress unique identifiers or where cell counts are too low? Is the data meaningful, and is there demand for it to be released? What are the financial costs involved (including burden costs to make it available).

DCMS has its own transparency site containing a number of orphan datasets. Data might have been released as an FOI response but it can quickly become out of date. A decision to release should consider whether the data will be updated and the frequency. If the costs to publish or release data are not judged to represent value for money, to what extent should the requestor be required to pay for public services data, and under what circumstances?

3. If the costs to publish or release data are not judged to represent value for money, to what extent should the requestor be required to pay for public services data, and under what circumstances?

A threshold needs to be applied, similar to £500 limits with FOI.

When a decision is being taken about whether to release data, the owners need to consider whether it is time-limited data, whether it is sustainable to keep the data up to date and whether the data is still relevant. The organisation should be clear if it is a one-off release (with explanations why), or document all future releases in their publication timetable.

4. How do we get the right balance in relation to the range of organisations (providers of public services) our policy proposals apply to? What threshold would be appropriate to determine the range of public services in scope and what key criteria should inform this?

This could cover all bodies under the statistics act, as these ALBs already have a good understanding of data quality issues, and releasing data to agreed standards (i.e. those outlined in the Code of Practice for Official Statistics). Potentially this could be extended to any organisation funded directly from public funds. We would like to see greater transparency in some private bodies which operate in areas controlled by Government.

(.....)REDACTED AT REQUEST OF DEPARTMENT.

5. What would be appropriate mechanisms to encourage or ensure publication of data by public service providers?

It is our view that policing this role could be quite difficult, and a body such as UK Statistics Authority would be best placed. Instead it could suffice if organisations list which datasets they would make available, and for these to be listed for all organisations in a central hub such as data.gov. Users should be allowed to comment and provide feedback on available data sources to improve usability or coverage, and these should be set out in a statistical work-plan, which is produced each year.

There has been a suggestion in ‘opening up commentaries’ as html releases. This may that run into objections from IT and Comms teams, as they have a very particular house style and further formats could clutter publication pages.

Policy Challenge Questions

1. **An enhanced right to data:** how do we establish stronger rights for individuals, businesses and other actors to obtain, use and re-use data from public service providers?

This could be established through legislation and parliamentary procedures.

2. **Setting transparency standards:** what would standards that support an enhanced right to data among public service providers look like?

The definitions would be much clearer, on exactly which data this would relate to, and what defines a “public service provider”. There also need to be limits to requests, similar to the £500 limits with FOI.

Metadata standards might be useful. These needn’t be all inclusive, like with the Dublin Core Metadata Initiative (DCMI) there could be a standard lexicon which could be used, where relevant, plus any custom fields. Across the repositories available to departments there are

different ways of describing the same data. It should be an aspiration to achieve a level of consistency across all the various data providers and publishers.

3. **Corporate and personal responsibility:** how would public service providers be held to account for delivering Open Data through a clear governance and leadership framework at political, organisational and individual level?

4. **Meaningful Open Data:** how should we ensure collection and publication of the most useful data, through an approach that enables public service providers to understand the value of the data they hold and helps the public at large know what data is collected?

Data should not be released for the sake of open data. It should have a purpose, be relevant and have a use in the public domain. User feedback should be used to develop and assist what is 'meaningful' data.

5. **Government sets the example:** in what ways could we make the internal workings of government and the public sector as open as possible?

The government could set up and encourage more open days, when government statisticians can showcase the work they have been doing. Journalists and developers could be invited to ask/work with the data owners. Government departments should also look to release more experimental statistics, but caveat accordingly with any necessary limitations. This could be published in statistics work plans on each government website, and could be located centrally on the ONS publication hub.

6. **Innovation with Open Data:** to what extent is there a role for government to stimulate enterprise and market making in the use of Open Data?

I don't think this should be a priority for government. We should be willing to work with and/or assist entrepreneurs but this is not something we should instigate. We also need to be careful we are not giving any given party a competitive advantage by working with them. How does the open data agenda ensure this does not happen? Many statistics are published with a caveat they cannot be used for commercial gain, and this would go against this proposal. The economic growth argument for open data appears to have been added so that it would gain backing from HM Treasury? The goals set out for this programme need to be clear, realistic and achievable.