

Making Open Data Real : A Public Consultation

Response by and on behalf of Landmark Information Group Ltd



Registered Office: 7 Abbey Court, Eagle Way, Sowton, Exeter, Devon, EX2 7HY
Registered Number 28928803 Registered in England and Wales

Website: www.landmark.co.uk

Author: Mike Williams, Principal Data Scientist

Date: Wednesday 26th October 2011

Version: Final

Consultation Response

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

A clear definition of “Public Task” is required, although where this is a statutory responsibility of an organisation this would be relatively clear where it is not this will be somewhat problematic. The latter could stray into areas where third parties already provide services using a particular data set or group of data sets.

Other organisations such as the EU and OPSI have found defining the “Public Task” challenging, any definition should not be open to interpretation or to avoid any protracted debates as to whether a particular data should be considered as “open” or not by virtue of the task creating it. A clear definition will also ease any data audits undertaken by any independent oversight organisations.

With any releases of data a consideration should also be given to the provision of suitable “metadata” which in part should form part of the definition of the data set. In certain instances this could take the form of a publically available specification.

The release should be free from implicit or explicit costs and have no proprietary dependencies.

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

The release of a data set should always be in the public interest and should follow an agreed “code of practice” which should be developed in partnership with the public and private sectors.

It is important that transparent terms exist for the release of any public data which attracts a charge between commercial organisations and those government agencies who receive the data as a part of a pan government agreement such as the PMSA.

Due consideration should be made of the private sector organisations who may already offer services around any data sets which currently attract a charge from the public sector data providers which may be candidates for being re-designated as “open”. Thus avoiding an potential negative affects to the existing markets.

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?*

Any tests that determine “value for money” need to be publically available from any bodies who govern the release of data. There should also be an option for the “requestor” to pay a premium for any such data on a cost recovery basis. However, any charges should be appropriate and respect the provisions of the Freedom of Act (FoIA) and the Environmental Information Regulations (EIR) legislation.

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?*

Any organisations that create data which falls within the agreed definition of “Open Data” as a part of their public services should fall within the scope of these proposals.

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

Encourage the further development of data.gov.uk in order to list any organisations which fall under the definition of organisations in point 4 above.

They should also be mandated to list any data sets which fall under the definition of “Open Data” on this site together with their ability to supply each data set that fall under the definition of “Open Data”. This could then act as a “virtual leader board” of performance.

Other mechanisms could also be considered in terms of making the release of “Open Data” a statutory requirement of the organisation, make provisions for withdraw of a certain level of funding from the organisation through targeted fines or impose the threat of audit so that the organisation needs to fully justify their decision.

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?

1) *How would we establish a stronger presumption in favour of publication than that which currently exists?*

There should be a presumption for the release of “Open Data”, this could be enforced through the development of statutory instruments where applicable. Other options could include making the release of data a pre-condition of receiving funding for services from the public purse or undertaking independent auditing of organisations.

2) *Is providing an independent body, such as the Information Commissioner, with enhanced powers and scope the most effective option for safeguarding a right to access and a right to data?*

It is important that such a body is established to provide governance in this area, options could include existing bodies such as the Information Commissioner, OPSI or the OFT etc, They must act in a transparency manner against an agreed code of practice. They should also have the remit to undertake independent audits of any organization in respect of the data which they release or more importantly don't.

The key to the success of any appointed independent body is that they have the appropriate powers to allow them to regulate in an effective manner to get the required results.

3) *Are existing safeguards to protect personal data and privacy measures adequate to regulate the Open Data agenda?*

We believe the current Data Protection Act and the associated powers of the Data Commissioner are sufficient to regulate the “Open Data” agenda.

4) *What might the resource implications of an enhanced right to data be for those bodies within its scope? How do we ensure that any additional burden is proportionate to this aim?*

No comment.

5) *How will we ensure that Open Data standards are embedded in new ICT contracts?*

Development of standard paragraphs to cover the provisions of open data together with a mandate to include them in new ICT contracts in order to release any associated public funding for such projects.

Setting transparency standards: what would standards that enforce this right to data among public service providers look like?

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

A prerequisite to achieving high compliance is the development and publishing of agreed definitions of “Open Data”, associated interoperability standards and an agreed code of practise. This will then provide a “yard stick” for organisations to be measures against.

It is important that a “bottom-up” approach is followed from the initial point of data collection until the point the data is published. At the point of delivery, such as via data.gov.uk, a rating scheme such as “Five Star Rating²⁰ for Open Data” could be employed to provide a quick view on the level of compliance. Any data uploaded such also have a minimum level of metadata, including source, format details, currency, geographical extent using such metadata standards as GEMINI 2.

However, any rating system should be readily understandable to all audiences as not all individual may be aware of World Wide Web Consortium open standards as RDF or SPARQL²¹.

Consideration should also be given to provide an easy feedback mechanism for the identification of data anomalies back to the data originator from which the data was originally published.

2) *Is there a role for government to establish consistent standards for collecting user experience across public services?*

No Comment.

3) *Should we consider a scheme for accreditation of information intermediaries, and if so how might that best work?*

We would suggest that there are already a number of accreditation standards that could be adopted by the “information intermediaries” such as those from BSI; Best Practice in Information Security Management & ISO 9001; Best Practice in Quality Management.

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?

- 1) *How would we ensure that public service providers in their day to day decision-making honour a commitment to open data, while respecting privacy and security considerations?***

Mandate the release of data as a part of their funding arrangements or as a part of their statutory duties and giving the consistent message of the presumption of release of data. Within a new transparent regime milestones could also be agreed or KPIs introduced in order for an organisation can be measured against their compliance to these.

- 2) *What could personal responsibility at Board-level do to ensure the right to data is being met include? Should the same person be responsible for ensuring that personal data is properly protected and that privacy issues are met?***

In order to ensure the “Right to Data”, one of the initial considerations must be does the it contain any” personal data” which does not fall within the remit of “Open Data”. Therefore, it should seem sensible that any such board members should have responsibility for both areas.

- 3) *Would we need to have a sanctions framework to enforce a right to data?***

If there is transparent publication of the details on each organization’s ability to supply their data sets on such a portal as data.gov.uk this should in part form a degree of sanction as this could act as some sort of “virtual leaderboard” of compliancy.

However, in certain instances that may not have the strength to enforce the supply of data. In such instances some sort of financial penalty could be introduced from the organizations funding body until the convention is addressed where the release is not a statutory requirement.

- 4) *What sectors would benefit from having a dedicated Sector Transparency Board?***

No comment.

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?

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

We would support the continued development of data.gov.uk as a central portal to discover what information is both available in an “electronic” format but also what information is held by an organization but not available in a suitable format for publication.

Although the supply of data in pdf format is a step forward it does not make the data contained in these documents to be readily usable. Other formats should be published in preference such as .csv, XML or JSON to allow the data to be readily consumed in databases and other information systems.

Any such portal should be consistent with metadata standards as adopted by the INSPIRE Directive or GEMINI 2.

2) *How should data be prioritized for inclusion in an inventory? How is value to be established?*

One option to prioritise the inclusion in an inventory could be on the basis of “demand” for a particular data set perhaps through the development of voting functionality on the data.gov.uk site. However, this does presume that individuals are aware of the existence of a data set in the first instance. Therefore, another approach could be for the public sector organizations to prioritize the release of their data perhaps on the basis on the amount spent creating the data set as a part of their public duties.

3) *In what areas would you expect government to collect and publish data routinely?*

Within the definitive core reference geographies in order to link the variety of themes of Open Data together. These include the Geodetic Framework, Geographic Names, Addresses, Streets, Land and Property Ownership, Statistical and Administrative Boundaries.

This subject is further explored in a paper by Professor Robert Barr - <http://www.agi.org.uk/storage/events/agi2009papers/RobertBarr.pdf>

4) *What data is collected ‘unnecessarily’? How should these datasets be identified? Should collection be stopped?*

Where a data set is not required to be collected as a part of the public duties of the organisation but is also collected by an alternative source then its collection by the public body should cease.

5) *Should the data that government releases always be of high quality? How do we define quality? To what extent should public service providers ‘polish’ the data they publish if at all?*

There should always be a goal of collecting quality data at the point of collection, however, it is acknowledged that this is not always the case. We would recommend that data should be released regardless of quality and acknowledged the status of the data as being “raw” within its accompanying meta-data. With greater visibility of such data feedback can be provided to the public service providers in order to identify areas that need to improve. Public service providers should not divert resources to “polish” any data they release where it is not required as a part of their function.

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

1) *How should government approach the release of existing data for policy and research purposes: should this be held in a central portal or held on departmental portals?*

We would suggest a central portal approach would be the most sensible rather than individual departmental portals. This would make the discovery of each data set far easier going to one search point, perhaps building on the existing capabilities of data.gov.uk.

2) *What factors should inform prioritisation of datasets for publication, at national, local or sector level?*

The publication of data sets should be independent of geographical scale and should occur when they become available. However, suitable metadata will also need to accompany any data releases so that data currency is clearly noted.

3) *What is more important: for government to prioritise publishing a broader set of data, or existing data at a more detailed level?*

The publication of data sets should be independent of level of detail and should occur when they become available. However, suitable metadata will also need to accompany any data releases which comments on the relationship of any data to other more or less detailed equivalents.

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?

1) *Is there a role for government to stimulate innovation in the use of Open Data? If so, what is the best way to achieve this?*

The primary focus of government should be the release of the data in the first instance and providing a framework of continual improvement for the release and quality of the data created. The role of innovation should be for others unless there is a specific government funded project that requires such an approach.