

## **Making Open Data Real: A Public Consultation**

A response to the public consultation by the Geospatial Engineering Panel (GEP) – a joint panel of the Institution of Civil Engineers and the Chartered Institution of Civil Engineering Surveyors.

### **1 Glossary of key terms**

#### ***1.1 Do the definitions of the key terms go far enough or too far?***

We have a concern about the use of the terms data and information. Throughout the consultation document when the term for data is used, it seems to imply information. This should not be the case; information and data are different. Data has to have some level of analysis, interpretation and/or manipulation applied to it to provide information.

We also feel that when data is provided that it should conform to the relevant **Metadata Standards** and be issued with associated metadata. This would provide some context to the data, its quality and hence the confidence that can be placed in it. If some form of redaction has taken place to the data to make the issuing of it more intelligible then this should also be communicated via the metadata.

Our concern is that the issuing and use of data without the appropriate metadata could be open to misinterpretation or misrepresentation which could lead to serious consequences.

There is a professional responsibility to the analysis, interpretation and/or manipulation of data and representing the output as information - the joint institutions would encourage that through sharing data the integrity of the professional is not undermined.

Some consideration should be given to the inclusion of "Public Task" in the definitions as this will be an important factor in deciding whether to make a dataset open or not.

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

There are many existing vehicles available that contribute to deciding whether to make a dataset open or not – Office of Public Sector Information (OPSI), Information Fair Trader Scheme (IFTS), Freedom of Information (FOI) , and Environmental Information Regulations 2004 (EIRs) to name but a few

It would be difficult to legislate a definitive set of data as “Open Data” due to potential personal and national security issues. Also the importance, relevance and hence value of any data cannot necessarily be made by the provider – it is ultimately determined by the user and/or the use it is put to.

For this reason the GEP would recommend that a procedure is established whereby some form of Public Task or business case could be submitted by the public/user and that this is arbitrated upon by an independent body – such as the Information Commissioner - or at least a body independent of government authorised to do so.

### ***1.3 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?***

The release of data in its original form (i.e. as held by the public service providers) should not be costly when the data is released in full and unaltered – i.e. in a recognised standard format, with associated metadata. Therefore in the majority of circumstances the requestor should not pay for public service data.

If data must be redacted or otherwise altered before release to make them into a single work or to provide context, then we would argue this is a necessary “internal” need anyway and therefore the costs should be borne by the holding organisation.

The judgement as to whether a dataset represents value for money lies with the user and the use they make of the data. The value is in the processing of data and the creation of meaningful information.

If a provider undertakes some form of data analysis or interpretation then this could be regarded as a value-added service and could be charged for. In this instance, consideration should be given to the final product. If it provided the original supplier of the data value, or another government body or public service value, then there should be a possible exemption. If however, the final product did not benefit a government body or public service then there should be a definite charge, comparable to the cost of analysing the data, the cost of providing a channel to service and the cost of delivering the data. However, in this instance, the public should still have the right to the raw data to undertake their own (value adding) analysis or interpretation.

### ***1.4 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?*

There should be an ultimate aim to make all government departments/organisations “in scope”. The Fol and Public Task will initially direct this.

It is very difficult for a public service provider to decide what data might add value and what data might not. The judgement as to whether a dataset represents value for money lies with the user and the use they make of the data. The value is in the processing of data and the creation of meaningful information.

For this reason the GEP would recommend that a procedure is established whereby some form of Public Task or business case could be submitted by the public/user and that this is arbitrated upon by an independent body.

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

If the public service provider or government body is to be measured against the publication of data, then the data will be published. If the public service provider or government body fails to meet the measure then penalties could be applied.

Currently in the delivery of services from a public service provider or government body, if there is a Key Performance Indicator applied against the service then it will be delivered.

The ICE and CICES are institutions which strive to promote the use of good quality data within industry, and so would be supportive in the publication of the good quality data

The publication of data by public service providers should therefore be mandated and enforced by bodies such as the National Audit Office (NAO) via the establishment of KPIs.

## **2 An Enhanced Right to Data**

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

Mandate it.

*2.2 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?*

We would encourage the introduction of an independent body, whether the Information Commissioner with enhanced powers would be the correct option is still open to debate.

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

Through the manipulation of a number of data sources personal data could be generated. By opening up the data exchange process, better management and regulation needs to be applied to the data, which could imply that the existing safeguards are not sufficient.

However it is clear that the Data Protection Act (DPA) should be the vehicle to regulate this.

*2.4 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?*

Initially an enhanced right to data may have resource implications for those bodies that fall within its scope. For example, to ensure that the appropriate levels of data security and standards are maintained some organisations may have to employ a data manager. There may also be charges associated with the redaction of data to prepare it for release etc.

These are no doubt potential additional costs but it could be argued that some of these functions should be in place anyway (best practise et al).

There could also be issues with hardware requirements and connection issues. The consultation is encouraging the use of “cloud computing” which we believe is commendable. However, the establishment and migration to this infrastructure may take time and will no doubt incur additional costs. Some of this cost could be justified by the increased benefits the release of this data brings whilst others could be negated by collaborative working and a joined up procurement process, which all government bodies and public service providers can sign up to. The Efficiency Review by Sir Philip Green highlighted several examples where a joined up procurement process by government could generate much saving.

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

A sound procurement process will ensure the delivery of acceptable Open Data standards. It is important that through good procurement the best value can be attained, along with sufficient scope to promote innovation.

In the same way that the Office of Government Commerce (OGC) improved government ICT procurement and contracts, the same approach could be applied to Open Data standards.

There must also be a clear link established in these contracts between existing legislative and/or regulatory obligations such as Freedom of Information (FOI) and Re-use of Public Sector Information Regulations 2005 (RPSI) respectively.

## **3 Setting Transparency Standards**

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

All data should be supported by Metadata which conforms to industry Metadata standards. This will ensure that all data supplied will meet the minimum requirements required for the data.

Ensure it is built into the audit process and the KPIs and ensure it forms a fundamental part of any future ICT contracts.

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

Standards should be led by providers and users. There are numerous example of government and industries working together to set standards and best practise – see British Standards and ISO for example.

The risk of the government establishing an “Open Data” standard is that one size does not fit all. Therefore, industry leaders and users should work with government to set the standards and also be tasked with promoting improvement within their field.

The ICE and CICES are institutions which strive to promote the use of good quality data within the Civil Engineering industry.

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

Yes, however, any scheme should be developed with industry. It is important that each industry is consulted as to its requirements. This in

turn will drive the need for accreditation and ultimately accountability for the data managed. For example this has worked well with regard to ISO standards and accreditation for the likes of Quality Management Standards.

## 4 Corporate and Personal Responsibility

### *4.1 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.*

Existing legislative and regulatory obligations cover this by way of the DPA, FOI etc. This would also need to be covered by some form of on-going information governance training.

### *4.2 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?*

It is important that the responsibility for data is held at board level. It has been shown that “personal responsibility” in the private sector has reduced health and safety issues for example, and so such responsibility should also be applied to data.

If “personal responsibility” at Board-level comes with a level of accountability and auditability that is measured and ultimately has consequences (such as loss of funding, support or, in extreme circumstances, loss of job) then this will ensure that personal data is properly protected and that privacy issues will be met.

### *4.3 Would we need to have a sanctions framework to enforce a right to data?*

Yes

### *4.4 What other sectors would benefit from having a dedicated Sector Transparency Board?*

Assuming that this sector, in the words of Francis Maude, covers “...every bit of government business.”, then other sectors to consider would be the utilities sector.

Despite the recent Upper Tribunal (Administrative Appeals Chamber) upholding the decision of the Information Commissioner and finding that privatised water and sewerage companies do not fall within the definition of public authorities for the purposes of the Environmental Information Regulations 2004, there are many people that feel this sector should have

much more transparency to data. This decision has already had significant ramifications with regard to public access to utility data.

## 5 Meaningful Open Data

### 5.1 *How should public services make use of data inventories?*

#### *What is the optimal way to develop and operate this?*

It is important that all data is recorded clearly using industry recognised metadata standards. This will ensure that all data can be quickly found and utilised appropriately. The inventory should also include all the caveats associated with the data, including accuracy issues and quality concerns.

The optimal way to operate and make use of data inventories is via the Office of Public Sector Information (OPSI) and the Regulations on the Re-use of Public Sector Information. These are already identified in the Key Terms in Section 1 and given that this is the existing regulatory requirement already available within the legislative framework – namely implementing Information Asset Registers (IAR) - surely this is already in hand. OPSI, operating from the National Archives Office, is already tasked with a wide range of services relating to “... finding, using, sharing and trading information.”

### 5.2 *How should data be prioritised for inclusion in an inventory?*

#### *How is value to be established?*

Most of these mechanisms for the prioritisation and/or inclusion of data in an inventory are already in place.

All data should be included, as data only provides value when it has been used to produce information. The outputs should then be stored in the inventory as data for future use.

### 5.3 *In what areas would you expect government to collect and publish data routinely?*

If data can be collected regularly then the data should be available. Only through the continuous collection, use and feedback can the overall quality of the dataset be improved, and the associated metadata will show the improvements over time.

It is appreciated that in some circumstances the regular and continued capture of data might be seen as an enhanced value-added (Premium) service. In this instance a cost recovery process would need to be in place to deliver this or the Taxpayer will have an increased burden.

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

The government departments and associated service providers are presumably collecting data for a purpose i.e. a justified business case has already been presented and approved. Therefore all data should be necessary. Through interpretation of the data and analysis of what use Public Data is put to should provide information on trends and use therefore justify continued collection (outside the original business case that was presented for its collection!).

#### 5.5 *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?*

Presumably the data being collected by the various government departments and associated service providers is “fit for purpose”. As long as it is published with accompanying and appropriate Metadata, then the user will be aware of the quality and will be able to accommodate and take account of the quality (poor or otherwise) into their analysis or interpretation accordingly.

It is appreciated that in many cases data may be inherited from a predecessor organisation and the current organisation cannot be held responsible for the quality of the data, although all data providers should be expected and encouraged to continue to strive to improve data quality where this is possible.

It should be appreciated that not all data will be of high quality – that on its own should not be a reason for a service provider not to publish it. GEP welcomes and supports the stance by Government to “...publish data of lower quality in preference to holding it back, while seeking over time to drive up the quality of that data.”

## 6 Government Sets the Example

### 6.1 *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?*

As long as there is a single version of the truth – i.e. the data held in one location - it does not matter. The purpose of a portal is to point you to that source. Either a central portal and/or a departmental portal could direct a user to where that data is located. After all, a portal is only a gateway or entrance to where that data is stored – there can be several gateways.



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

All data will have a level of appropriateness, which would be reflected by why and how it has been collected. If data is to be published then it should be accompanied by the appropriate Metadata standards. The prioritisation of datasets to be published should be demand driven wherever possible.

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

Again this needs to be demand driven.

## 7 Innovation with Open Data

*7.1 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?*

We believe that government should work with industry through its Technical Advisers Groups (TAGs) to stimulate innovation in the use of Open Data. Industry and the general public will have clear ideas as to how they want to use the data, and so should be involved in the innovation of the use of the data. Public interest and commercial opportunity will be more significant stimulators to innovation than any role government can invent. On the whole government has neither the skill, imagination nor the culture to innovate. Where there is a demand and opportunity to innovate government should support this by ease of access and legislation.