

**BCS, The Chartered Institute for IT  
Consultation Response to:**

**Cabinet Office: Making Open Data Real: A Public Consultation  
Dated: 27 October 2011**

**BCS**  
The Chartered Institute for IT  
First Floor, Block D  
North Star House  
North Star Avenue  
Swindon SN2 1FA

This page is left deliberately blank.

## **BCS, The Chartered Institute for IT**

The Institute promotes wider social and economic progress through the advancement of information technology science and practice. We bring together industry, academics, practitioners and government to share knowledge, promote new thinking, inform the design of new curricula, shape public policy and inform the public.

As the professional membership and accreditation body for IT, we serve over 70,000 members including practitioners, businesses, academics and students, in the UK and internationally. We deliver a range of professional development tools for practitioners and employees.

A leading IT qualification body, we offer a range of widely recognised professional and end-user qualifications.

[www.bcs.org](http://www.bcs.org)

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

**Dated: 27 October 2011**

### **Consultation Document:**

<http://www.cabinetoffice.gov.uk/resource-library/making-open-data-real-public-consultation>

## Consultation Questions:

### Glossary of key terms

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

The Institute believes that the terms require greater definition and expansion. The terms appear to focus on static data and not transient data or information flows, for example, streaming data. There is also GIS data, planning data, time series based data (that allows cause and effect modelling) and data used for evidence for example video. These need to be accounted for.

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

The Institute considers it important to refer to all discussions and debates undertaken via the Office of Public Sector Information (OPSI) in the implementation of RPSI and the former Department for Constitutional Affairs in implementing the FOI, before a decision is made.

It is crucial that the data is provided in a format which is accessible to most forms of analytical software, both commercial and open source.

It would also be important to establish who would apply the tests. The UK Statistics Authority, the Office of National Statistics (ONS) and The Royal Statistical Society (RSS) should routinely be involved, in our opinion, when revision of tests is under consideration.

#### 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 Cabinet Office should revisit existing debate and research undertaken by the OPSI. This would save time and resources.

There will be some commercial organisations which will want to use public data. If they are to make profit from the use of such data then it is fair that they should pay a reasonable charge. However, if non-commercial organisations (especially voluntary) or individuals use the data then a charging structure would undermine the whole purpose of Open Data. The system should contain a clear definition of who should be charged.

Without a default position, it is difficult for voluntary bodies or private individuals to challenge, on evidential grounds, conclusions which have been drawn by well-funded groups (e.g. developers) or by tax payer funded groups (e.g. local authorities). Voluntary bodies frequently have greater experience and expertise in data analysis than such funded groups. Making data open to challenge can only improve government. This is why charging would in general be counter-productive.

There needs to be clarification as to who decides whether or not the data is considered “value for money”. The Institute believes provision should also be made to inform those requestors of similar requirements so they can join together and improve value for cost.

#### 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?

The Institute believes that this should be implemented like the FOI, which applies to companies that are wholly owned by public authorities.

It is important to acknowledge that certain industry sectors through statute already have to provide information to assist governments e.g. gas emissions, environmental health information and occupational health information.

At the moment it is impossible to define the scope. The Cabinet Office should start with a presumption that any public data should be open and that there should be a list of carefully worded exclusions from that default position.

In order to strike a balance, the Institute believes that it would first be important to establish a framework of different types of Open Data use. This would also help assist in the governance of data and thus identify the priority of data.

Such a framework could consider:

- i) Reference data that can be used by all agencies e.g. British UK addresses, all roads in the UK, talent and skills in the UK by location.
- ii) Citizenship information that can be used by several departments and agencies e.g. healthcare statistics reports for the NHS.
- iii) Commercial cost information that can be used to identify commercial opportunities e.g. the number of claims against a local council for inadequate road maintenance based against road maintenance contracts for the same road versus the safety of that stretch of road against maintenance of the road.

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

The Institute would consider making Open Data a legal obligation. We would argue that there should be two procedures, one for requests to individual bodies, and a second one for what are effectively information gathering exercises.

Mechanisms should include:

- Mandating supported by the National Audit Office remit to assess maturity of compliance with the requirements.
- Financial incentives for data publication.
- Financial incentives for the quality of the data produced.
- Financial incentives for frequency and update of data.
- Ability to collect revenues from the digitisation of historic data.
- Corporate responsibility for data openness.
- Shares or profit sharing from the use of data by commercial organisations.

## An Enhanced Right to Data

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

The Institute considers that mandating would provide the solution and would force the right to Open Data to become embedded as part of 'business as usual'.

We further consider that:

- There should be a clear cross government set of data classification standards which cover structured, unstructured and other media types.
- There needs to be a clear separation between information that has been 'aggregated' or 'summarised' and information that would allow 'business insight' i.e. almost 'raw data'.

- All government agencies should publish the type of data and its usage.
- The minimum associated information that has to be provided should be identified and associated with a dataset so that there is consistent information being provided e.g. how often the data is provided/updated, how changes will be incorporated, how to request additional services e.g. data currently being provided once a year may be valuable and could command a levy if it were provided monthly, new data being introduced and change in delivery format.
- Data formats should be published and data standards for certain types of data provided.
- There need to be clear guidelines on:
  - Archiving of Open Data.
  - Retention of Open Data in easily accessible information.
  - Ability to request data to a defined period of age.

## 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?

The Institute considers the Information Commissioner to be a sensible arbiter, who should be given a broader scope in responding to concerns about privacy, efficacy of data and disclosure rights.

With regard to the devolved administrations, we do not agree that it is up to such administrations to determine which of the data should be open, or even which data should be gathered. Comparable data must be gathered from each of the constituent administrations in the UK. That does not preclude an administration gathering data of particular interest to it, but it does mean that all the administrations should gather certain agreed data. Such agreement should be very detailed, down to the definitions of variables and units of measurement.

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

We believe that the Data Protection Act provides sufficient safeguards in protecting personal data and privacy measures. It is envisaged that different data cells will be linked and therefore consideration will need to be given to how these sets of data can be used in conjunction.

## 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?

The Institute considers this to be complex to calculate, as there are a number of factors:

- Public available data may have to be provided in a one off 'electronic format' with details of certain parameters and updates.
- Public data may need to have an explanation of the different data fields so that information can be:
  - Aggregated.
  - Linked to other open data.
  - Linked to other data.
- Where internal systems are unique identifiers for certain types of information, e.g. National Insurance number and/or NHS number, there may be a need for anonymising information.

The 'right to data' at the moment is probably not an effective costed item or reported item across different sectors. The use of management time, information communication and technology resources and administrative personnel to respond to requests or clarifications would be better served via self service mechanisms where a data consumer can access the data. This in itself would drive business case rationale for cost effective mechanisms to share data and collaborate with different agencies.

The data provided needs to take into account the requirements from organisations on the value of data to their organisation e.g. today we may publish reports and analysis once a year requiring an organisation to double key data and provide commentary throughout the year based on other data sources.

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

The Institute believes that this can best be achieved by ensuring that there is a clear link between the Open Data standards and the existing legislative or regulatory obligations i.e. the FOI and RPSI. The legal departments need to therefore be included. The procurement departments need to be informed of what new contract terms need to be embedded in all contracts from the earliest possible opportunity, just as the introduction of the FOI did. The Office of Government Commerce (OGC) provided significant guidance in this area and needs to be updated to reflect the Open Data requirements.

There are a number of other considerations:

### Quality of Data

Quality of data is a challenge as the industry and user groups will demand different levels of quality of data and frequency of data for their use. Setting an initial minimum may be the more appropriate standard and having an approach that allows for assessment and data quality improvement through time/value. Quality of data may have a commercial value and policy for data needs to consider how come companies may obtain data at a different level of quality as the level of quality data may assist those companies in differentiating their services commercially.

### Intellectual Property

Data model ownership needs to be established and the provision of Open Data needs to be in models that are both 'open' and do not infringe copyright and intellectual property rights.

### Cost of ICT services

Procurement is impacted as COTS based products may not have appropriate interfaces and security data services that meet Open Data standards.

COTS products may have to develop specific interfaces that may increase the costs to central government.

The storage of this data may increase the overall costs of data.

### Consistency of data

New systems need to take into account access to Open Data and the transformation and change to Open Data into new formats.

## Setting Open Data standards

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

The Institute believes that standards that allow for usability and interoperability need to be combined into the Data Quality Audit that is carried out by the Audit Commission. Best practice standards from The Chartered Institute of Public Finance and Accountancy (CIPFA), Audit Commission, Socitm and the Local Government Association (LGA) should also be embedded.



We believe that Open Data requires a data management and lifecycle approach. This will require close assessment of the data. A customer address for DEFRA for a farm is also a commercial business address for HMRC, is a sequence of land registry points for land registry and is a personal and commercial address for TV licensing and the DWP.

Furthermore, the Government's interoperability infrastructure layer needs to be expanded to cover utility services, composite applications and data integration.

There is no mention in the principles of the need for Data Quality i.e. with regard to the data being released. Although the need for an accurate inventory is mentioned there is no mention of data quality and this therefore needs consideration

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

The Institute believes that the government should play a role and that it should consider advice and guidance that has already been published. The table below, for instance, is derived from available public sector sources.

### The six characteristics of quality data

Characteristic	Description
<b>Accuracy</b>	Data should be sufficiently accurate to present a fair picture of performance and enable decision-making at all appropriate levels. The need for accuracy must be balanced with the costs and effort of collection. A prerequisite is that definitions for data should be specific and unambiguous. The data must be at an appropriate level of detail to influence related management decisions, and must be within a reasonable margin of error (the amount of variation normally expected from a given data collection process)
<b>Validity</b>	Data should clearly and appropriately represent the intended result. Where proxy data is used, organisations must consider how well this data measures the intended result.
<b>Reliability</b>	Data should reflect stable and consistent data collection processes and analysis methods across collection points and over time, whether using manual or computer based systems or a combination. Managers and stakeholders should be confident that progress towards targets reflects real changes rather than variations in data collection methods.
<b>Timeliness</b>	Data must be available for the intended use within a reasonable time period. Data must be available frequently enough to influence the appropriate level of management decisions and should be current enough when they become available: for example, it may be appropriate to accept a small degree of accuracy where timeliness is important.
<b>Relevance</b>	The data reported should compromise the specific items of interest only. Sometimes definitions for data need to be modified to reflect changing circumstances in services and practices, to ensure that only relevant data of value to users is collected, analysed and used.
<b>Completeness</b>	All the relevant data has been recorded. Monitoring missing or invalid

	fields in a database can provide an indication of data quality and can also point to problems in the recording of certain data items.
--	---

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

We believe that a scheme for accreditation of information intermediaries should be considered. This is especially true for those companies deemed to have Open Data that can be provided by public services, then the provision of the data would be required.

## Corporate and personal responsibility

### 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.

We believe this to be ongoing and needs to be embedded into the culture of management as part of 'business as usual'.

We feel that the Cabinet Office should also give consideration to the following:

- Auditing of all new systems and processes to ensure alliance to open data policy.
- Review and adjudication of any issues raised about Open Data.
- Make Open Data part of the framework of government including inclusion in specific organisational structure decision making bodies and inclusion within broad management job specifications.
- The role of the Chief Data Officer for larger organisations.

### 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?

Personal responsibility needs to come with a level of accountability and audibility that is measured scored against and thus has consequences.

More specifically, we believe that the Cabinet Office should consider:

- Corporate reports illustrating the value that Open Data held by the public service body has provided to other organisations.
- Measures for the response times for information requests.
- Measures shown for the number of "self-service" information requests have been provided.
- Progress to agree initiatives to prove access to data is shown, including quality of data.
- Public/open discussion is carried out for new provision of data.
- Records are shown for where public data has been found to be misused.
- Clear penalties for data misuse are laid out.

### 3. Would we need to have a sanctions framework to enforce a right to data?

Yes, there is a clear need for leadership and the mandating of standards. Further, if Open Data is to succeed then people must be held to account to achieve stated aims and objectives.

### 4. What other sectors would benefit from having a dedicated Sector Transparency Board?

We consider that the energy and utilities sectors would benefit from a dedicated Sector Transparency Board.

The Cabinet Office should also extend the question to incorporate policy areas that should have a Transparency Board e.g. environmental policy that covers all industries.

## Meaningful Open Data

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

The Institute believes that a link needs to be made between the existing Office of Public Sector Information and the Regulations on the re-use of Public Sector Information, as identified in the key terms in section 1. The Information Asset Register (IAR) already incorporates data inventories into the legislative framework. Focus should therefore concentrate on the implementation of IAR.

As part of any new system or refurbishment of an existing system, either manual or computerised, data provision and repository of data should look at the existing inventories and determine the re-use of the data.

Data inventories also need to consider Master Data Management practices and how they may provide multiple “access points” to data e.g. mirror sites where the data can also be accessed should a source of data not be available.

### 2. How should data be prioritised for inclusion in an inventory? How is value to be established?

The Institutes believes that inclusion should be based on the consideration of the following:

- Number of users using the data e.g. number of public services that use the data.
- Frequency of use/demand number of times used, even if data is replicated elsewhere.
- Frequency of demand or querying of such information.
- To provide fact based evidence, informing the populace of reality and reducing uncertainty.

We would again argue that it is worth referring to all previous discussion and debates undertaken through OPSI in seeking to implement RPSI, in order save time and resources.

### 3. In what areas would you expect government to collect and publish data routinely?

We consider the following set of information should be routinely published:

- Administration of government.
- Schools, town planning, universities, health organisations.
- GDP information.

### 4. What data is collected “unnecessarily”? How should these datasets be identified? Should collection be stopped?

There are existing examples of vast amounts of personal data that are being collected on old forms that are never used. Form filling needs to be reviewed across all organisations with the intention of justifying the requirement to collect, use and store the data. Also, duplication of data entry between government departments needs to be reviewed.

For example, a highly qualified and trained security specialist has security clearance at the highest level from the Cabinet Office and yet the Metropolitan Police requires the individual to fill in their own security clearance forms, which will ultimately be sent to the same central location for them to look up

the same information, putting it on a different form. We would question whether the resources could be better used elsewhere.

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?

Quality of data should be appropriate to its purpose. Provision of data if inaccurate will be challenged by users.

Data quality criteria should be stated for different classifications of data. Through time the polishing, accuracy and /or the enrichment of data should be undertaken by the data owner i.e. the public body that has been given primary responsibility for the data and its collection, maintenance and in time deletion.

Where data is changed and translated then this may move away from the underlying factual data. When there is a variation in information between open data and ‘polished’ data, both interpretation of the data and polishing of the data need to be clearly understood.

### Government sets the example

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?

The approach should be simple and transparent. There should be an appropriate symbol that is published on all websites that will lead them to a portal of available datasets.

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

The Institute believes that prioritisation should be determined on need/requirement as presented by those wishing to use the dataset.

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

We consider that this will be driven by market demand in different areas at different time e.g. UCAS results and placements have a yearly cycle.

### Innovation with 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?

Given the current economic difficulties, the Institute believes that the government priority should be restoring economic growth. We further believe that the private sector is well placed in stimulating innovation.

We consider there to be great examples of Open Data in action contained in the document. These need to be shared widely in order to provide tangible evidence of what is possible and the benefits of doing so.

**End**