

Making Open Data Real

Intellect Submission

October 2011

Introduction

Intellect is the UK trade association for the IT, telecoms and electronics industries. Its members account for over 80% of these markets and include blue-chip multinationals as well as early stage technology companies. These industries together generate around 10% of UK GDP and 15% of UK trade. Intellect provides a singular voice for these industries across all market sectors, and is a vital source of knowledge and expertise on all aspects of the technology industry.

Submission

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

1. Intellect feels that this initiative would benefit from further definitions. For example, there are references made within this document to the 'knowledge economy' and yet there is no explicit definition of 'knowledge.'
2. The 'dataset' definition refers to data as being a bi-product of service delivery. Intellect feels there needs to be a mechanism for Government to understand who is relying on this data, which would affect the wider definition of 'dataset.'
3. Intellect member companies feel that the scope of 'data' is limited when placed in the context of the numerous companies involved in the data corporation.

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

4. Given the rationale of 'Open Data', data should be published unless there is a specific reason not to do so.
5. However, it should be clearer who is making the decision about the publication of single data sets.
6. In addition, Government has a responsibility either to make sure the published data is accurate or to indicate how accurate they believe it to be.

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?

7. The majority of Intellect member companies are of the view that the minimal costs to departments that the publication of data implies are worthwhile if this overall policy is of benefit to the UK economy and society as a whole. However, some Intellect member companies feel that the publication of data may affect established markets that are currently providing services around information publication.
8. There is a risk that as more data is made available publicly at no cost, the Government may start to suggest that responsibility for analysis of data is incumbent upon others, including citizens. Despite its availability, many citizens will not be able to interpret raw data effectively meaning the Government risks misinterpretation if it does not provide both the raw data and suitable description and analysis of the data (meta data).

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?

9. Many Intellect members believe the scope should be for the entire public sector including local government and NHS but excluding the Single Intelligence Account and Protectively Marked Material.

10. There is, however, a risk that some information not viewed as personal, could be inferred from other datasets to achieve a picture of individual citizens which unintentionally exposes personal information.
11. Intellect feels strongly that new EU data protection legislation due in November 2011 will impact the organisations included in the scope of these policy proposals significantly.

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

12. The most appropriate mechanism would be the mandating of a presumption of publication of this data. The initial aim should be to rapidly publish the available raw data, with migration towards data being published as 5-star linked open data.

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

13. The most effective means of achieving this would be to extend the scope of the National Archives to seek and encourage publication on behalf of the Government.

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?

14. As each individual has rights in relations to their personal data and amending them, the ICO should be empowered to ensure this happens; these enhanced powers should be accompanied by greater reporting responsibility for the ICO. The ICO should enforce both the “pull” and “push” of open government data in a more transparent and accessible way for citizens.

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

15. Generally, yes. It would be sensible to address this question once the changes to the Data Protection Act have been implemented in November 2011.

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?

16. There are clearly advantages and disadvantages to the provisioning of more open data. The right to data is likely to create more work for bodies within this scope. A clear communications campaign on the impact of frivolous requests should be conducted to minimise these incidents. Bodies should see to put in place adequate process and resource capability to manage this supported by guidelines or shared services from the ICO.

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

17. This should be appropriately addressed in the terms and conditions of contracts and frameworks. Contract specific terms and conditions should be permitted to add flexibility to the application of a common ‘Open Data’ clause for ICT contracts.

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

18. The Open Knowledge Foundation has developed the Open Knowledge Definition (OKD), endorsed by Sir Tim Berners-Lee, which can form the basis for high common standards.
19. The Cabinet Office has recently run a survey on Open Standards and the outputs from this should be used to inform any decision.

20. Intellect member companies feel that there should not be an undue focus on technical or data format standards; the principle underlying open government data licencing is that they should be content or data that is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and share-alike.

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

21. Yes. Some appropriate standards already exist in this area:
- a. Tim Berners-Lee suggested a 5-star deployment scheme for Linked Open Data which has global adoption. Its scope and focus is on interoperability and data quality.
 - b. Existing international standards on how citizens log on and interact with their respective governments can be found at www.data.eu, which rates European governments for their digital interface with citizens.

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

22. This concept of accreditation of information intermediaries is not appropriate for open data.

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.

23. Intellect feels that there should be an assumption that all data should be made publicly available unless there is an explicit statement made as to why they are not, to over-ride this principle. This presumption of publication should be incorporated into departmental business plans to ensure effective implementation.
24. Intellect believes that Government should review the lessons learned from the implementation of the Freedom of Information Act when seeking to drive a commitment to open data throughout public service providers.

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?

25. While Intellect feels there should undoubtedly be a Board member that is responsible for protecting personal data, it is important that elaborate governance structures are not implemented to enforce this. It would be appropriate for the individual responsible for information management and FOI requests to have their remit expanded to include enabling the right to data and presumption of publication.
26. We recognise that the organisation, rather than an individual, may be legally accountable unless the organisation is operating in line with the Companies Act and all applicable legislation.

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

27. Intellect believes that a sanctions framework would over-complicate this situation unnecessarily. With the Information Commissioner and the Data Protection Act in existence, there is no need for further legislative regulation.

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

28. It would depend of the scope of 'dedicated Sector Transparency Board' and the extend to which is might introduce red tape (See response 27).

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

29. This importance of understanding what data is available is enshrined in point 8.13 of the consultation document that states '*a right to data is meaningless without knowledge of what is available.*'
30. Intellect feels that there is convergence around CKAN, the open-source data portal software. CKAN makes it easy to publish, share and find data. It provides a powerful database for cataloguing and storing datasets, with an intuitive web front-end and API. This mechanism allows any citizen access data inventories. While this is a useful tool, it demands a certain level of expertise to use the data itself.

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

31. As the previous question suggests, there are a plethora of data inventories. Value can only be comparable if these inventories are interoperable and federated.

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

32. Intellect feels that all non-personal data of this type should be made available unless there is a clear and stated reason not to do so, which would be subject to challenge under the right to data via the ICO. However, we feel that the delivery of that collection and publication is a service that can be commissioned from the private and third sectors. As such, Government should release as much data as possible, but collect as little as possible where other organisations already do so.

What data is collected unnecessarily? How should these datasets be identified? Should collection be stopped?

33. Intellect feels that as a general rule of utility should be applied to the publication of datasets. If the data is not put to use by the public body, other organisations or individuals, then it should be viewed as unnecessary. This is clearly extremely hard to gauge before the data has been published.

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?

34. It is the view of Intellect member companies that if the quality of the published data can be high then this is preferable, but the principal concern should be to publish the available data.
35. If it is possible to determine the accuracy of this data then the Government should indicate their assessment of the data's accuracy when published.

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?

36. Intellect members feel that the location of the data is not important once the data are published as other organisations will point to the data of relevance to their constituents.
37. A central portal is unnecessary if data inventories are federated.
38. There remain doubts as to whether the Office for National Statistics data currently published only for academic research purposes should be for public consumption. This is essentially a question of licensing, for which the National Archives is responsible in this area. The National Archives and Information Commissioner will need the Charity Commission to make explicit their definition of not-for-profit organisations, as these should include Social

Enterprise Mark certified organisations and other bodies working in the public interest for whom policy and research data is useful, not just Registered Charities.

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

39. As stated above, the assumption should be that all data be made available for publication unless there is a clear and stated reason not to do so.

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

40. The overriding principle should be for the data to be published as quickly as possible.

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?

41. Intellect members feel that it should be Government's role to publish data and the private sector should drive innovation around this resource.

42. Where successful initiatives have sought to stimulate innovation, private capital has been key to this process, notably with projects such as the Cabinet Office Innovation Launchpad.