

Open Data Consultation

Submission from the Locus Association

October 2011

1. Introduction

The Locus Association welcomes the chance to contribute to this consultation on Open Data. Should the Transparency Board require any further information or clarification on any of the points raised then please let us know.

Locus would like to take this opportunity to underline its commitment to working with the Government, executive agencies, other Public Sector Information (PSI) holders and regulatory bodies towards the maintenance and development of a vibrant, information-driven UK economy that ultimately works to the benefit of the public sector, private sector and the end consumer.

2. The Locus Association

Locus was established in Autumn 2005 and formally launched in January 2006. The Association's President is a former Director General of the OFT and OfTel, Sir Bryan Carsberg.

The Association exists to raise awareness of, and promote the development of, a healthy and competitive private sector in relation to Public Sector Information (PSI), including PSI. Locus acts as a forum for exchange of information, keeps its members up to date with latest policy developments, and provides advice and guidance. Our members, which include ESRI, Dun & Bradstreet, TeleAtlas, Landmark, and eMapsite, recognise the potential of PSI, but also the benefit of uniting to address some of the challenges and pooling resources to meet them.

3. Open Data Consultation

In his introduction to this Consultation, Francis Maude MP correctly observes that "the best way to tap into the UK's tradition of creativity and invention is to give that [PSI] data away". The data, in effect, already belong to citizens. The State should not attempt to either deny them access to data unreasonably or use its monopoly position to trade in the data.

Where government acts commercially, distortions occur and the result is sub-optimal for the citizen as well as the public and private sector. Several reports commissioned by government and at least two select committees have come to broadly the same conclusions on these points.

Locus members believe the Consultation describes the opportunity and benefits well. It may well be that the benefits will be substantially greater than those envisaged.

The challenges include: to implement the Open Data policy in a way that does not disappoint expectations of what the State can reasonably deliver at a time of national stringency, to continue to protect the privacy of individuals and keep their confidence, and to redirect existing public sector information-collation activity along more productive lines for the future.

These issues are complex and risk becoming more so unless there is a real attempt to impose simple principles. This paper proposes that Government should focus on doing only what it needs to do for "good government" – and that it reappraises how it arrives at these decisions following the Canadian example. "Good government" would therefore include the principle of releasing data ("open by default") but would not include enhancing or spending money to achieve its release except in particular circumstances (e.g., social gain – subject to independent review). Where release for re-use requires further cost beyond that which government would expect to provide for its own purposes then the private sector should be asked to pay this additional cost – in effect a charge relating to the marginal cost of distribution. Locus members support this approach.

1. Consultation Responses

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

Whatever solution is adopted it is likely that more thought will need to be applied to definitions over time, especially that of "Public Task". A key success factor will be the extent to which Open Data is used as a springboard by the private sector to develop value-added services.

If the private sector believes or suspects that the public sector might wish to expand its own activities to offer services potentially or actually available from the private sector — with or without payment — then the private sector will naturally remain reluctant to invest. Thus the definition of Public Task becomes a critical issue.

Where Public Task relates to Statutory responsibility it is straightforward but this only covers some activities.

Work done recently by OPSI concluded that Public Task could be more widely attributed to "custom and practice". This is a vague and adjustable notion. OPSI's briefing paper also assumes that the public body concerned should define its own public task, a less than impartial approach. The EU is similarly struggling with definitions of Public Task.

Current UK proposals assume some mechanism will be available to appeal against a statement of public task. This appeal process must have the confidence of private sector re-users or it will be ineffectual.

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

The consultation makes a number of reasonable suggestions. All PSI should be Open Data "by default" unless there is a reason that its release might cause significant damage to third parties. As suggested, a code of practice should be introduced.



There should be a right of appeal to the Information Commissioner whose powers should be enhanced.

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?

This question leads to another: who decides what is "value for money" and how can Government ensure that the holder of the data isn't abusing its position by requiring payment? Perhaps the clearest way forward would be for OPSI to set general rules related to the type of data requested (for example, is it "data" that can be extracted quickly and easily or is there a value-added/aggregation/cleaning process required to produce it?), with standard prices for activities beyond simple extraction and provision of data.

Locus members generally do not object to paying reasonable charges for data. Today, however, it is often difficult to predict what a public body will seek to charge for the provision of data, if it will release that data at all.

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 data that satisfies the "Open Data" definition should be in scope in principle. However government in general finds itself in an inherent conflict given the level of outsourcing, PPP and PFI involved in execution of public sector policy objectives.

While such relationships are competed for on a level playing field, they will be subject to all manner of negotiations in the final terms and the subsequent execution of the contract. Similar tensions exist in housing associations, educational and hospital trusts and all manner of areas where "subsidy" in its myriad forms plays a part in the operations of the public sector framework (transport, defence, emergency services, justice, private sector R&D and so on).

This is a far more complex area than it might initially appear. Arguably what is required is the establishment of metrics and transparency requirements for each sector or PSIH or combination thereof and which apply to PSIH or executing entity at all times, forming if necessary statutory and contractual reporting protocols.

The bottom line is Government and the regulators must apply far more stringent data publishing requirements on PSIHs of all shapes, sizes, and areas of responsibility.

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

It should become a statutory requirement for all government entities that they publish/release all data that falls under the suggested definition of "Open Data".

A Public Data Corporation could also create best practice and other guidance with case studies and standards references that would help accelerate this process.



Questions for consultation: An enhanced right to data

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

It should become a statutory requirement for all government entities that they publish/release all data that falls under the suggested definition of "Open Data". Training and the creation of KPIs within PSIHs would also help.

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?

Those that wield the powers, be it ICO, OPSI, OFT, TNA or other authorities need to be knowledge and practice leaders in this area, providing guidance, case studies, exemplars, support, templates and the like as well as an 'accessible' sounding board (perhaps incorporating PDTB) long before formal enforcement measures are required.

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

Locus members believe existing safeguards are likely not adequate. There have been a number of incidences of "lost" and "stolen" laptops and discs from Government employees which would indicate that more needs to be done. Kieron O'Hara's report and the plentiful technical research in this area, particularly around deanonymisation, suggests that there are many serious challenges. It seems increasingly necessary that Government understand the risks to privacy that deanonymisation represents and introduces protocols to the appropriate regulations to increase protection for personal data in the face of a commitment to otherwise open data.

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?

It's difficult for us to respond to this; it will depend on the existing resources of each individual PSIH.

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

Government has found it difficult to agree and then implement common standards in the past for reasons which are unclear but may relate to inadequate pressure being applied from politicians and Department Heads. Having common Open Data standards is essential to the success of the policy.



Simple, transparent steps could include:

- define the acceptable standards, protocols and approaches;
- mandate them in terms of service, contracts and elsewhere with a PSIH specific series of milestones;
- disqualify tenderers who refuse to comply;
- discipline personnel who fail;
- publish milestones and performance;
- penalise PSIHs for performance against the agreed milestones for adoption.
-

Questions for consultation: Setting Open Data Standards

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

To start with, by agreeing and publishing simple steps:

- Agree standards, particularly for inter-operability
- Publish them
- Agree milestones with PSIHs
- Create accessible resource base (templates, guidance, use cases, best practice, cost benefit analyses, etc.)
- Publish metrics regarding success in achieving milestones
- Penalise persistent failure to achieve milestones.

Department Heads could be made responsible for releasing appropriate data under the policy and agreeing that all appropriate new data will adhere to Open Data standards.

Government needs to sharpen its analysis of the cost and benefit of its data creation and maintenance, how it might best be procured and to what standards — and significantly improve its skills in these areas.

In the future it is quite probable that PSI may increasingly include data obtained from the private sector. Where private sector data is included, such as Royal Mail postcodes or from Google maps, then it is vital that the government has licensing arrangements in place that enable it to distribute the data under the Open Data initiative or valuable public sector datasets will become inaccessible.

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

There may be a role for Government to create a simple framework for the collection of information on different types of users and their experiences. However, this should probably not be an immediate priority.

3. Should we consider a scheme for accreditation of information intermediaries, and if



so how might that best work?

There are already Government entities that have accreditation systems, from Trading Funds such as Ordnance Survey and the Met Office to Executive Agencies such as the Environment Agency to research organisations such as the British Geological Survey or even CLG's Planning Portal. It could be useful to attempt to standardise and streamline these existing systems to the extent possible, rather than to create an entirely new scheme.

Questions for consultation: 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.

See above – mandate, publish, support, measure and, ultimately, penalise.

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?

Public sector “boards” are few and far between outside those who have a trading activity. However, each PSIH should appoint an SRO or equivalent and all PSIH staff need to have within their terms and conditions a requirement to meet the above mandate as directed by the SRO. The SRO should be responsible for agreeing and then meeting milestones and should ultimately be accountable for both success and failure in delivering the open data vision for that PSIH.

Protecting personal data is an increasingly technical challenge and the responsibility to ensure that that technical/analytical responsibility is fulfilled (through internal resourcing or external appointment) should fall under the remit of the SRO.

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

As above, yes, although the stick and carrot of agreeing milestones and having the success in achieving them measurable and measured and published is likely to be good enough except for the most intransigent PSIH.

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



The question suggests that a whole series of quangos be established harnessing sectoral expert knowledge. This seems an unnecessary and costly overhead for the open data agenda to have to accommodate. If the responsibilities and requirements are mandated and the milestones agreed on a PSIH by PSIH basis then there is a framework for open data release that is auditable, transparent and ultimately subject to penalty. So the direct answer to the question is no; however, establishing a PSIH specific programme for implementation very much equates to a harnessing of PSIH data knowledge in pursuit of the wider goal.

Questions for consultation: Meaningful Open Data

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

Locus members believe PSIHs would benefit significantly in terms of the direction and speed of travel towards an improved open data environment if they were provided with as many tools and as much support as is possible to release their data. Therefore frameworks with common language, consistent terms, examples, generic guidance and best practice would be useful. data.gov.uk provides a good a portal as any for data publishing. The quest for Government is to establish an appropriate entity that can deliver these frameworks; the Public Data Corporation offers one option while equipping an existing entity such as ICO or TNA with the appropriate resources and skills might offer another and demonstrate the UK government's commitment to the open data agenda.

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

Within the scope of the Open Data policy, Government should focus on its own needs and not those of the private sector. Priority and value should be priority for and value to good government, not to third parties.

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

Where data is required by Government for its own purposes; and where data stimulates good government and the ability of third parties to hold government to account. Otherwise, Government should not be in the business of collecting and publishing data it does not require for its own purposes.

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

If data is not necessary for good government then its collection should be stopped forthwith. The data audit mentioned above might be used to help uncover such anomalies.

There exists a strong suspicion (and some hard evidence) that some of the data collected by Trading Funds could be collected and managed by the private sector more efficiently.



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?

It is unrealistic to expect all government data to be of high quality but, at the same time, it is clearly unacceptable that public funds are spent collecting a sub-optimal dataset ("If you cannot measure, you cannot manage").

However, whatever the antecedents of the data, it should be published where it falls within the Open Data policy and no attempt should be made to improve its quality unless the private sector agrees to pay for the cost of doing so (e.g., Property Intelligence plc with the Valuation Office Rating List).

Questions for consultation: 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?

Government should be seen as "just" another end user of open data and should be able to avail of it through the outlets adopted by PSIHs and Government more widely, notably data.gov.uk. It should discourage the creation of PSIH-specific machine-readable feeds and APIs where possible, and encourage wider use of data.gov.uk.

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

Government has no role to play in second guessing what data sets should be prioritised for release as open data. If the data passes the tests as open it should be released. If it is necessary to somehow decide which data should be released first, this should be done based on which data is requested by potential users.

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

Government policy should be to make all appropriate data Open and accessible. Where this requires investment then the user must be asked to provide it. Users are likely to be the arbiters of what level of detail they would require.

Questions for consultation:

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?



Government should not be in the business of picking winning companies, models or technologies. Unlike 'research', innovation implies or invites the notion of direct commercial involvement carrying with it inference of seed/grant funding that sustain start-ups and other in the absence of a business model, market or other business fundamental.

Public service providers should focus on their core remits and should not in any way "promote the use of data" as long as they have fulfilled their mandate to publish the open data that they generate or are otherwise responsible for. Anything more would give public service providers with a value-adding role that takes them beyond the collection that data. This would not only add costs but also establish new tensions around where value lies and whose open data priorities should be addressed first, or at all. The level playing field that is satisfied by releasing the data covered by the tests proposed will mitigate the extent to which this can happen.

