

Open Data Consultation: Questions

The following section sets out a series of challenging questions in six key areas to start the debate.

The six areas identified as presenting opportunities for change, creating both demand for data and supporting and enabling supply, are:

This is a response from Keith Murray.

See also attached “background notes” which I have also submitted to the PDC Consultation.

1. **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?**

Right to data: governments should act as “enablers” to stimulate the resources of the nation be they public, private third sectors and individuals. In this case they have the power to manage resources at their disposal. Public data is a massive and often untapped resource. Some data is available but with so many constraints – third parties often find it easier to go out and collect it again – others make their data available because they see themselves as enablers.

Hence the public data principles, which have not really been communicated well, should act as the starting point for policy that applies to *all non-personal public sector data*. The onus should be to opt out – where grounds are proven that it would not be in the national interest to publish a dataset eg security. This may require legislation of some kind to support its adoption. [See EU INSPIRE Directive – without the legislation the current drive to improve interoperability and adopt common standards would have run into the sand some years ago].

However the implementation of this is not as simple as it sounds. There is already a lot of duplication of data in the public sector (see background notes – even in single organisations). There is also a lot of poor quality data that would be difficult to use. It is questionable that simply putting up a lot of unusable data is going to help anyone – especially when we are seeking reuse to save costs.

What we really need are discrete definitive data objects published in a way that we can selectively connect them and use them out of the box. For example with the Highway Agency data – the value lies in traffic accident data, data flows, road surface condition surveys – we don’t really need a multitude of disjoint road network datasets published.

Two decades of piecemeal developments cannot be resolved overnight – in which case there will need to be a set of options open to those organisations obliged to open up their data. These would allow fast track adoption and those that would adopt within a

timespan that allows time to migrate to a reuse model.

2. **Setting transparency standards: what would standards that enforce this right to data among public authorities look like?**

Transparency “standards” is strange term – but INSPIRE has effectively covered this for location data by establishing a set of 34 themes of data and supporting it with legislation (Framework Directive & Regulations). It is not a standard per se but uses standards as part of its implementation framework ie ISO, OpenGeospatial stds. Note most of the detail is in the Technical Guidance – which is non-binding. This is a nice balance in that the Regulation forces a minimal change but it would be a fool who expended time and resources in an implementation that did not follow the Technical Guidance.

While it could require significant work to address all other data domains in Britain beyond location & INSPIRE – it is conceivable that different domains could develop their own protocols – for example Statistics people already do this.

Those domains would require some over-arching principles to start from – eg the current or an enhanced set of the public data principles – the five star path to publication would be an excellent starting point.

3. **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?**

For some time there has been a concept of x-government coordination groups that take on the accountability for delivery of their data domain. The Location Council is one candidate that is attempting to thread a way through this terrain – largely in isolation at the moment. We could see a number of like bodies for statistics, environment, transport, health, finance, education. Indeed many of these already exist in some form – though the future should be more radically different from their current agendas.

4. **Meaningful Open Data: how should we ensure collection and publication of the most useful data, through an approach enabling public service providers to understand the value of the data they hold and helps the public at large know what data is collected?**

As noted in my PDC submission – simply dumping poorly thought out and badly engineered data on a stall is not going to help anyone apart from encourage a cottage industry trying to improve n different versions of it.

Data has to be made easy to use “plug n play” – hence API’s have to be part of the story

in terms of reuse and scalable adoption by many.

Overall of course and also as noted in the PDC submission the primary vehicle going forward to really support reuse and interface x-organisational datasets is the publication of data in Linked Data form. By publishing common reference data in this way and encouraging the public sector to use it in situ (and not down load it and build yet another silo) – we then have a powerful foundation to build all kinds of true value add services.

This need not cost significantly – it may just be about spending “Business As Usual” funds in a different way – see the Environment Agency Bathing Water linked data pilot.

Very quickly people will discover these sources - but what they really need is some of assurance that this network is sustainable (like some of the better websites today such as the BBC) – ie the next policy decision is not going to reverse the decisions.

Data.go.uk has made a great start – but it’s only a means to an end – the end is the application > where much of the underlying technology is completely hidden – the application simply does what it claims to do – eg: <http://legislation.gov.uk>

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

Relating perhaps to the earlier governance points – any coordination of policy and implementation should clearly engage the wider user. It can be a challenge to find someone representative of a user community in a coordination group as the Location Council discovered – but supported by dissemination of papers and records this largely works.

Community engagement in terms of communication, setting priorities, progress, prototyping and testing are all essential activities to establish engagement and early adoption as well as feedback.

6. **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?**

Government has to demonstrate leadership – but it has been doing that well over the last couple of years. Because an application often requires data from many sources – these may not come together quickly –especially as there is not the same pressure on everyone to release and modernise data immediately.

There may be several goals:

- do things more effectively – cut operational costs in the process
- provide better public services – easier to use and engage with at all levels of

society – business to citizen

- Innovation in services – public and private or in partnership.

There is scope with open data to do much under innovation once more datasets come on line - supporting the citizen, business user and consumer.