The National Endowment for Science, Technology and the Arts (NESTA)

CONSULTATION RESPONSE

Making Open Data Real: A Public Consultation

NESTA is the UK's foremost independent expert on how innovation can solve some of the country's major economic and social challenges. Its work is enabled by an endowment, funded by the National Lottery, and it operates at no cost to the taxpayer. NESTA is a world leader in its field and carries out its work through a blend of experimental programmes, analytical research and investment in early stage companies. For further information, visit http://www.nesta.org.uk

Through our work on various programmes, most notably 'Make it Local'[1], NESTA has developed insight into the power of open data and the potentially enormous benefits it can deliver to the UK. While these are well articulated in the consultation document it is NESTA's view that the potential benefits in opening data are not confined to improvements in transparency and accountability in the public sector (as may have been the previous consensus) – in fact, open data can provide the biggest gains by enabling the development of useful services . Before this potential can be unlocked however there are some significant challenges that need to be overcome. These include endeavouring to open more data to the public, ensuring this data is relevant, useable, dynamic and linked, ensuring an emphasis on data produced at a local level, and making sure that those seeking to use the data either have the skills necessary to do so or are connected to those who have. One area where there is need for much attention is in the generation of evidence to assist those attempting to formulate open data policy. Particularly, in relation to this consultation we find that there is insufficient evidence available to provide substantive answers to some of the questions posed. In light of this, we have addressed some of the overarching questions below by focusing on some principles relating to open data that we believe need testing. This response proposes some ways that evidence or the conditions for generating evidence might be created.

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

[1] NESTA - Make it Local, http://www.nesta.org.uk/areas_of_work/public_services_lab/make_it_local

There is a need for the terminology to be clearer in relation to the public ownership of online services, i.e. open data is quite separate from other forms of IP that gets created when online public services are commissioned. For example, there may be intellectual property which is created (design, software code) that does not constitute public data, but which is suitable for commercial exploitation. Being clear about what information/IP is owned will assist the commissioners of online services to identify what data can be made open for public use.

A large amount of open public data is not in a format that allows the "semantic web" to make best use of it. Therefore, we recommend that open public data is published in a format that allows it to be "Linked".[2]

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

It is NESTA's belief that all newly-generated data should be "Open by Default", with the onus being on the collector of the data to provide evidence as to why they believe the data should be kept closed. While testing some issues such as whether or not opening data compromises individuals' rights to privacy is likely to be feasible, testing for other factors such as the cost versus benefit of opening particular data may not be so. The costs of opening data depend on a range of factors such as the current and future costs of securely storing data that is not made public and the effort in ensuring data reaches standards set across a particular sector. The benefits of opening data may be impossible to calculate given the plethora of uses that may be made of a given dataset.

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?

Currently we find that there is a dearth of evidence that could be used to answer the above questions. One way of creating such useful evidence would be the trialling of different business models to examine which models can effectively serve the area. We feel it would be extremely difficult to propose how much should be charged for data and in what circumstances without evidence on how such scenarios would play out in a market setting. A suspension of costs should be imposed as it would allow different models to be tested to give data providers and those attempting to establishing policy in the area a more informed view of how and when to charge for data. Locking the data away or asserting prices will only serve to reduce the experimentation in this area. Once the value and costs of providing different datasets has been tested, it would be possible to start charging for updates following a 'Freemium' business model. Mechanisms such as auctions allowing potential new service developers to bid for commercial licences may in time be another useful way of assessing market value, but only when there are published data about the value created by existing services.

Having said all of this, we acknowledge the cost implications of making a large number of datasets freely available. A less ambitious option is to open a number of datasets that have strong potential to spawn useful third-party services, and pilot their up-take and business models in the way we

[2] http://en.wikipedia.org/wiki/Linked_Data

suggest above. This would also make it possible to develop and test different standards and infrastructures for open data that could then rolled out more widely across the public sector.

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?

In principle the range of bodies obliged to publish their public data might extend beyond just the providers of public services and local and national governments to include all publicly funded activities. This may not be feasible in all cases, for example when private organisations receive only a minimal amount of public backing, and perhaps when funding is given through a publicly backed intermediary but even in such cases an approach could be taken to require publicly funded projects to open the data they produce from these activities. However, the 'net' of who should be obliged to make their data open should be cast as wide as possible.

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

Again this is an area where the generation of more evidence would be extremely helpful. There are many potential mechanisms that could be tested in this regard. One is a challenge model where public service providers are rewarded based on how they fare on for example the use of the data they produce, the usefulness of their data, how well linked their data is etc.

Another methodology is to encourage and incentivise data providers to work more closely with those that can transform the data into useful information. NESTA's 'Make it Local' programme is an example of how this can happen. The project showed how local authorities can work with digital agencies to unlock their data and provide really useful web-based services for their citizens.