

Fischer Family Trust.

Questions for consultation (Page 6)

The Government would welcome views on the following:

1. Do the definitions of the key terms go far enough or too far?
2. Where a decision is being taken about whether to make a dataset open, what tests should be applied?

Relevant criteria would include (i) In the public interest (ii) Would not harm individuals or organisations (iii) Would not allow individuals to be easily identifiable (iv) Supports improvement for services or individuals and (v) Cost Effective

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?

Does this imply that the provider of the data needs to make a judgement about the potential impact of how the requestor might use the data against the cost of providing the data in order to assess value for money? If so, this would seem to be a difficult task for the provider to undertake. It might be better to require providers to classify information into 4 areas (a) freely available (b) available but at a defined cost (c) available but with cost which will be determined at the time of request and (d) not available.

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?

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

The starting point should be a 'presumption of publication' along with clear criteria which would enable providers to test whether data should NOT be made available and potential users to understand the basis on which any decision not to release data had been made.

Questions for consultation (Page 25)

The Government would welcome views on the following:

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

In principle, yes – depending, of course, upon the powers and scope of the Information Commissioner's role.

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

Safeguards should be strengthened in terms of the publication of data where the number of individuals in the group is small. In terms of education data, it is common practice to restrict access where the number of pupils falls below 10. We do not believe that this is defined by regulation and feel that, in the context of open data, that regulations should include such a requirement.

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?

- *Key fundamental issue – open data and transparency is generally considered to be a 'good concept' but could potentially have a 'perverse' negative impact on vulnerable groups and*

government policy. i.e. whilst a key component of government policy is to improve life chances for those 'disadvantaged' pupils and families (e.g. introduction of pupil premium), those who may benefit least (at least relatively) from increased public access to data may be those very same 'disadvantaged' families. There is a risk that the release, interpretation and manipulation of greater amounts of data could actually increase the gap between the 'information rich' and 'information poor' in 3 key areas: access to data, understanding and the ability to act on data. This isn't to say that more data shouldn't be released to the general public (and obviously it can play a key part in improving life chances for disadvantaged groups), but both the Government and the 'market' should look to address some of the issues highlighted if possible (e.g. improved 'public' web access, clear guidance on use of data, support systems/bodies (e.g. schools, connexions, CAB, voluntary/NfP, private sector) to help 'users' to both understand and act on data)

- *Simply providing access to data will not provide economic and social benefits – the key will be to understand, interpret this data along with the ability to act upon it!*
- *Providing adequate context (either background information or built into data – e.g. value added or contextual value added data in schools) is important. For example, whilst providing 'raw health outcomes' data at hospital or consultant level may be seen as useful, it is also open to misinterpretation, especially by 'non-experts'. What data is provided, how it is shown and the guidance that goes along with it is key.*
- *Although data would be made freely available to all is there a danger of data being 'monopolised' by larger organisations that have the capacity to monopolise this 'data market' and use it simply as a 'profit tool'?*

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

Questions for consultation (Page 28)

The Government would welcome views on the following:

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

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

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

A process of making data more widely available should also bring with it a responsibility to ensure that the data is being used appropriately. This could be achieved through a combination of defined standards and accreditation.

For example, a common issue with interpretation (and misinterpretation) of data relates to the issue of confidence intervals. Standards for publication of data could require users of data to indicate confidence intervals when they publish data.

A code of practice could set out the standards to be applied when using or publishing data made available by public bodies. Users of the data would then, as part of any accreditation process, be required to accept the code of practice. Should users of the data be found not to be following the code of practice then the information commissioner should have the power to remove their right to use the data.

Questions for consultation (Page 30)

The Government would welcome views on the following:

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.

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?
3. Would we need to have a sanctions framework to enforce a right to data?
4. What other sectors would benefit from having a dedicated Sector Transparency Board?
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?

In principle, data released to a wider audience should be of high quality.

An example here is in the publication of data on school performance. Unamended data is available to schools and local authorities in October and there is a process whereby schools feed back updates and corrections before amended data is published in January. For some schools the differences can be substantial and it could be misleading to allow wider access to unamended data.

However, analyses provided to schools and local authorities using unamended data are valued by users and we would not wish to see a situation where the introduction of open data led to a requirement that unamended data could not be used.

We would suggest that a useful distinction could be made in terms of the intended use and ‘audience’ for the data. If the user of the data is providing analyses and data to those who would know and understand its status and reliability (i.e. schools and local authorities in this instance) then providing analyses based upon unamended data would be allowed. Provision of analyses and data to a wider audience should not be allowed until amended data is available.

Questions for consultation (Page 31)

The Government would welcome views on the following:

1. How should public services make use of data inventories? What is the optimal way to develop and operate this?
2. How should data be prioritised for inclusion in an inventory? How is value to be established?
3. In what areas would you expect government to collect and publish data routinely?
4. What data is collected “unnecessarily” ? How should these datasets be identified? Should collection be stopped?

Questions for consultation (Page 33)

The Government would welcome views on the following:

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?

Provided that data can be linked together the distinction should be irrelevant – modern ICT systems should be able to provide a user with a ‘unified’ interface and access to data even if the data is distributed across a range of systems.

A central portal providing access to all data could become very complex in terms of users finding the data which they require – it would almost certainly require some sort of ‘sub division’ – so departmental portals which are designed to common standards and can be linked together in a central portal as required might provide the most flexible solution.

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

Value for money – which would include an assessment of the cost of making the data available and the likely impact upon services and improvement. In order not to be a complex and time-consuming exercise this would need to be a ‘broad brush’ judgement made in relation to each dataset against defined criteria.

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

Initially, publication of the range of existing data which could be made available together with an assessment of whether it is in a form that can be linked together (or not) would enable an informed discussion of this question.

Additional Comments

A key concern – and one which is difficult to define – is that of the need for users of the data to act responsibly.

Perhaps an example might illustrate the point. The Fischer Family Trust (FFT) Data Analysis Project provides, as part of a wide range of analyses, estimates to enable schools to think about targets for improvement.

When providing analyses for self-evaluation, FFT have always provided two models of analysis: value-added and contextual value-added. However, because of concerns regarding the potential to reinforce patterns of under-achievement, FFT have refused – even when requested by users – to provide forward-looking estimates based upon the contextual value-added model.

We feel that providing data to a wide range of users (all schools and local authorities in England and Wales) brings with it a responsibility to try to ensure that data is provided in a manner which encourages positive aspirations and improvement and is presented in a manner which tries to guard against potential misuse or misunderstanding.

The consultation document refers to high levels of ICT literacy. Whilst there may be some questions as to the accuracy of that statement we feel that a much bigger issue is that of data literacy.

Whilst it is harder to define areas such as this we feel that, in enabling wider access to data, Government has a responsibility to try to ensure that users of the data act in a responsible manner – and this should include a requirement that users of the data take account of ‘data literacy’ issues in terms of how they use data, how they present it to others and in the decision they make about what to provide. Government also has a responsibility to put in place requirements and processes which would improve levels of data literacy, thus reducing the dangers of inappropriate or inaccurate use of the data likely to be made available under the Open Data process.

*Dr. Michael Treadaway,
Director of Research,
Fischer Family Trust.*