

‘Making Open Data Real’: Consultation Response

Charles D. Raab
Professor Emeritus and Honorary Professorial Fellow
School of Social and Political Science
University of Edinburgh
28 October 2011

1. I welcome the Government’s Open Data initiative because I think it holds the promise of changing the UK public sector significantly from one of information secrecy or difficulty of access to one of greater transparency and accountability. It also could also have a considerable impact on local and other communities in enabling people to participate and engage with public issues on the basis of better information available, in principle, to a wide range of participants. A great deal depends on how Open Data is implemented. In this there will be a good deal of bootstrapping, in the sense that important elements of cultural and organisational change, including better data standards and usability, must precede or at least run parallel to the open availability of data that, in turn, will promote further cultural and organisational change through transparency and accountability.

2. This consultation response will not comment on all the many issues and items in the consultation document. Nonetheless, a crucial element in Open Data – one on which I would otherwise have offered comments – is not dealt with in the consultation paper: the question of the relationship between public access to data and the protection of information privacy. The Minister’s Foreword pledges that privacy will be protected, and that the question will be considered in greater detail, including anonymisation and pseudonymisation. This is encouraging, although the timing of the consultation document did not allow consideration to be given to the commissioned report on Privacy and Transparency subsequently published by Dr. Kieron O’Hara (to which I contributed some thinking). It is to be hoped that a subsequent paper from the Minister will give the O’Hara report’s analysis and recommendations full airing, commentary, and possible further development towards practical integration with the main lines of Open Data as set forward in the current consultation document.

3. For the moment, therefore, I refrain from giving my response to O’Hara’s work because this was not called for in the consultation. However, there is one important point that needs highlighting because the consultation paper might have been unwise to say, in 6.1.3, that ‘at no point in this document does the data we describe refer to personal data’. This may or may not be true, for – as O’Hara discusses at length, what counts as ‘personal data’ is far from easily determined, even where such data has been considered to be anonymous because of the application of certain procedures. This question is currently at the heart of a great deal of debate and thinking in academic and practical circles, and it is likely that the quoted passage above would have had to be reconsidered if the O’Hara paper had been able to be reflected in the document.

4. Moreover, the sentence at 6.1.3 does not square with with other parts of the document. For example, in 7.5, reference is made to individuals’ access to their personal health records; 4.2 says that Open Data ‘is [inter alia] about giving people access to their individual medical records’; and one of the three ‘significant benefits’ of Open Data claimed in 5.1 is the ‘Right of individuals to access and control their

own service user records e.g. a school or personal health record'. These statements are inconsistent with 6.1.3, but it is more important to note that the right of subject access to one's own data already exist under the Data Protection Act 1998 (as they did under its 1984 predecessor), and therefore the claim of Open Data to be bringing it about is rather odd. I note that 5.2 states that this right of access 'is not considered further in depth in this paper'; all the more reason to regret the awkward non-coincidence of timing of the document and the O'Hara report. In addition, I note that among the questions for consultation under 8.7 is one concerning data protection in the context of Open Data. I think it is not appropriate to ask this without having a fuller discussion of privacy protection/data protection in this document, but it would be extremely important to ask it in a further consultation that took the O'Hara report on board. It is puzzling that this appears in the current document given the disclaimers about the discussion of the issues concerning personal data in the anticipation of O'Hara's report. I have therefore refrained from commenting on this question, but would be very interested in doing so subsequently if there is a further consultation opportunity following the current one.

5. Turning to other parts of the document, the Policy Challenge Questions (section 8) seem appropriate. More specifically, I would answer some of the questions posed in 8.6 as follows:

An openness default: this would be useful, and could help to shift the culture in the direction of the presumption of openness rather than secrecy. This is what many open government initiatives in the past failed to do, although FoIA has helped.

Proactive publication: this too would shift the balance, and is to some extent echoed in the publication-scheme requirements of FoIA but needs strengthening. How this requirement – and indeed, much of the Open Data agenda – would articulate with FoIA needs further explanation.

A powerful independent body and a right of challenge: this would be desirable in giving teeth to the initiative, although the final clause in its present form, about the rarity of over-ruling local management, could weaken this provision.

Information Commissioner's powers: It is right that the Information Commissioner should play a central role in Open Data. The exercise of his powers to the extent required depends in considerable part on the resources available to ICO. Public trust in Open Data as well as its efficacy would be harmed if the Government did not provide ample means for the ICO to act, and to bring about an increase in his powers beyond the recent extensions. The latter have largely been on the data protection side, which is not the focus of this consultation document. But insofar as this document has mentioned individual rights of access to personal data, it would seem possible to broach the question of the ICO's lack of powers to enter and inspect the data processing of non-public sector bodies under the DPA; and if such bodies play a part in the provision of public services (see the Glossary, page 5 and the Open Public Services White Paper), it would follow that any lack of powers of the ICO to enter and inspect these bodies in regard to their Open Data/FoI activities would be equally anomalous and unsatisfactory. However, it is not clear to me whether the involvement of non-public sector bodies in service provision always occurs through statute, or instead through other means, and therefore the Glossary definition might be restrictive.

Time limits for internal reviews: The ICO's recommendations seem sensible, and raise the game in a situation where delays are unacceptable.

Procurement rules and 'Open by Default': This seems a good step to take, particularly in considering ways of improving the situation in future. Procurement requirements could be a powerful way of engineering change in the direction of Open Data by removing an important source of frustration.

6. The proposed measures for practical implementation and standards, discussed in 8.9, look like they are going in the right direction, especially if linked with the corporate responsibility mechanisms proposed in 8.12. A Code of Practice formalising Public Data Principles would establish valuable expectations and send an important signal. Ways of publicly using the star ratings to reward or praise public agencies could be explored through the compliance monitoring process. The possibility of citizens' challenge is also important. Perhaps the establishment of citizens' watchdog panels, whether by locality or by public-body sectors, could help to prevent these accountability measures from being subverted or neglected. Such bodies would play a different role from the national Local Public Data Panel and Sector Transparency Boards, but could be articulated with these.

7. The fundamental importance of standardisation, data quality and metadata is recognised in 8.10, and the involvement of GSS will be crucial. In addition, it might also be valuable to involve the academic statistical and public administration community in shaping this work, perhaps through their professional and learned societies.

8. Making Open Data part of information governance and corporate responsibility, as indicated in 8.12, is necessary if Open Data is to be taken seriously at all levels and not relegated to organisational and technical back rooms (as data protection has tended to be for a long time). The Caldicott Guardian model has not always worked as well as it should, according to a review conducted some years ago, but it could be adapted for Open Data. In terms of corporate responsibility, staff training and organisational culture change, it would be relevant to see what lessons could be drawn from the practical experience and results of the changes made following the Data Handling Review, established several years ago in the aftermath of breaches of personal data by government departments.

9. The suggestion (page 31, top and 8.15) that only 'meaningful' data should be collected and published, and that data that has no value should cease to be collected, could be dangerous if it feeds suspicion that there are political reasons behind the termination of certain statistical series. In this light, the episode involving the Rayner Review of official statistics, many years ago under the Thatcher Government, is a cautionary story that needs re-telling. That said, the questions asked on pages 31 and 32 must be addressed, probably through wide consultation amongst interested parties and the general public, with subsequent proposals put out to consultation. However, it is obvious that this process could be interminable and subject to special pleading, and with some interests under-represented, although the Open Data machinery now in place could steer the decision-making on these crucial questions.

10. Government could set an example by undertaking the suggestions in 8.17, and should actively involve the external research community in finding better ways of using data for research purposes through its Research Councils, professional

UNCLASSIFIED

associations and learned societies. I have no immediate answers to the three more specific questions posed on pages 33 and 34, or to the one posed on page 36.

UNCLASSIFIED