Response to Consultation: 'Making Open Data Real: a Public Consultation'

Please note: this preamble should be read in conjunction with our responses to HM Government's formal questions - which follow.

1. Preamble

1.1 The Association of Census Distributors (ACD) was originally set up in 1993 as the Association of Census Agencies. The name was changed post the 2001 Census, when the term Census Agency was dropped by ONS.

1.2. The original purpose in setting up the ACD was to provide a forum for its members to negotiate with OPCS, the fore-runner of ONS as the collector of the national Census of Population. Subsequently the ACD also represented its members in discussions with other data providers, such as Ordnance Survey (OS) and Royal Mail (RM).

Current members of the ACD are:

Acxiom UK
Beacon Dodsworth
CACI Information Solutions
Callcredit Marketing Ltd.
Experian
PB Business Insight (MapInfo)

1.3. ACD members have first-hand experience of the value of Open Data, which goes back to the release of small-area data from the 2001 Census in 2003; these data were free at point of use, funded by ONS' Census Access programme.

'Free' Census data had a huge impact on the market, bringing the detailed (neighbourhood-level) data within reach of all commercial users for the first time. During the ten-year currency of the 1991 Census data, the OPCS royalty regime was such that a client requiring GB-wide data at Enumeration District (ED) level (some 125,000 areas averaging 175 households each) would need to pay some £100,000 for 100 variables, putting this dataset out of the reach of all but the largest organisations.. Post - 2003, the roughly equivalent data (actually available at Output Area (OA) level, 230,000 areas averaging 108 households) was virtually free of charge.

So, targeting solutions improved hugely and the industry flourished; to the benefit of end-users, distributors, and the government, which enjoyed the resultant tax revenues. The number of companies producing neighbourhood classifications increased from 6 to 11; the free data acted as a spur to new entrants, and to innovation.

1.4. Paradoxically, the one fly in the ointment in this scenario was Ordnance Survey (OS). ONS had used some OS data in the production of OA boundaries; ONS paid OS a

one-off sum in consideration of OS' IPR, and ONS made OA boundaries free to use (and royalty-free) for ten years, for all users. However, OS specifically excluded commercial resupply from the agreement with ONS.

OS brought OA boundary data within the scope of its Framework Agreement, although this dataset was clearly not an OS product (OS did not produce it, and did not even have access to it). Thus, if an ACD member should want to distribute OA boundaries, together with the Census data that populated these boundaries, that ACD member would be obliged to sign up to OS' Framework Agreement, plus a Specific Use Agreement - both of which were very long, very detailed, and very onerous on the signee, involving royalties payable to OS.

The ACD's view was that this did not add up. Census Access meant that Census data were distributed free of royalty to all end-users; given this, how could it make sense for commercial resuppliers to be subject to a royalty on OA boundaries, particularly when ACD members' end-user clients could get these boundaries free directly from ONS ? It seemed to us that this was simply not 'joined-up', and placed an impediment on the free dissemination of Census data.

1.5. This case was put to OPSI in October 2006, and was countered by OS. OPSI, while being sympathetic to the ACD's case, found that OS were not breaking PSI regulations, or IFTS regulations; but commented that OS was 'undertaking a fundamental review of its licences at present'. In the event, this review was not completed until after the DCLG consultation on OS in early 2010 had been completed. Given the introduction of OS Free, we understand that OA boundaries post the 2001 Census will be unencumbered by OS royalties.

1.6. The reason for mentioning OS in the context of IPR, pricing and terms & conditions, is that another area of concern for the ACD is that of addressing data. We have followed the founding of GeoPlace with interest, but were concerned to hear that, while the public sector will receive the National Address Gazetteer (NAG) data free of charge, the private sector will have to source NAG from OS. We gather that the expectation is that the pricing for NAG 'is likely to be in line with OS' other addressing products'; if true, then sales of NAG to the commercial sector are likely to be minimal, in our view. OS' terms & conditions will exacerbate this situation. We noted APPSI's Executive Summary of its response to last year's OS consultation stated that 'OS should not have any Intellectual Property Rights in derived data'. This advice has not been followed. The fact that OS' IPR runs through products derived from AddressPoint is a strong disincentive to its use by developers, and we assume the same scenario will apply to NAG.

1.7. It is clear to us that a single, definitive national address gazetteer is highly desirable; a two-tier pricing strategy would miss the opportunity, with ensuing disbenefits to the economy. We were heartened to read in the Transparency Board's June 2011 minutes that 'The Board set out their view that a single national address file, created by the merging of the GeoPlace data and the Postcode Address File needed to be available as a public good and as core data infrastructure'.

We agree entirely ! We believe that NAG should be Open Data, because that way, efficiency can be maximised, innovation and enterprise can be encouraged, and we will once again see the sort of rewards that flowed from Census data being made Open Data in the last decade.

In summary, we are in favour of Open Data - it works.

2. Specific areas of interest to ACD members

As an organisation, the ACD has quite specific areas of interest. So we propose to concentrate on the questions that relate to those areas of interest. We have no expertise in some of the other areas covered by the consultation document; therefore, it would add little for the consultation team if we attempt answers to them.

Specific areas of interest to ACD members are:

2.1 Opportunities to access existing data sources on better terms, ideally as Open Data (e.g. the National address Gazetteer, NAG)

2.2. Identifying additional data sources that are relevant to our activities, for enhancing existing models or services, or identifying new opportunities.

2.3. Pursuing opportunities to improve the overall data supply marketplace.

Association of Census Distributors - Member Companies
Acxiom
Beacon Dodsworth
CACI Information Solutions
CallCredit Marketing Ltd.
Experian
PB Business Insight (MapInfo)

Consultation Questions.Follow …

Response to Consultation

Glossary of key terms

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

The definitions are fine.

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

How useful it is likely to be to how many users
The government has argued for transparency. With this in mind putting data sets out that have already been created and seeing what users find useful would be a good way to achieve this. Deciding what to include/exclude would be time consuming and might prejudice the users by what the government thinks is useful. It could also lead to excuses for not releasing really useful data sets. Users do not know what is on offer so let them choose by releasing all available data sets.

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?

They should be asked to pay actual costs incurred

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?

N/A

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

Official instructions from the top.

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?

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

    Official instructions from the top.

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?

    Strong regulation is clearly necessary

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

    We believe so.

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?

    N/A

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

    Vigilance by those thus tasked.

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

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

    N/A

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

    Set up a User Forum, or similar

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

    Sounds complicated and time-consuming; but could have some value.

4 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?

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?

Official insructions from the top

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?

The Caldicott approach, referenced on P.29, sounds a very sensible way of handling this.

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

This need to be set up formally, with a Regulator in charge.

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

N/A

3. 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?

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

Have a central portal (www.data.gov.uk) which takes users to appropriate department for the relevant data. Have a good search and metadata to provide good information to help people access these data that they want.

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

Value should be established by establishing the degree of interest (i.e., likely volume of users) for the dataset in question.

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

Social statistics, address list (singular !), housing data, etc.
All areas which provide data for the government to work, .e.g. crime, health, benefits, head count estimates, immigration, inequalities, expenditure, census.

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

N/A

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?

Ideally of high quality, but there is often a trade-off between getting it quickly, and getting it absolutely correct. The ACD favours the former; with metadata describing data quality and any deficiencies.

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

    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?

    It does not matter where, so long as there is clear signposting of what is there and how to get at it.

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

    An estimate of the likely demand for each dataset.

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

    In general, the ACD would prefer existing data at a more detailed level - down to the lowest level that can be safely published. But some indication of what else could be made available would be very useful; easily published on a website, grouped by topic, and with email alerts to all those expressing an interest.

5. 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?

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?

The more potentially useful data that is made Open Data, the better. We regard the unlocking of such data as a priority for government; prioritised in the way outlined before.

Put the data up on the relevant department's website with good metadata and an example of how to illustrate these data, e.g. interact report or map. Explain how to use these data.