

Open Data Consultation – Making Open Data Real

Hackney Council Response

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

There are no definitions of Records in the glossary of key terms- this could be important in terms of defining information and ultimately datasets as part of the information lifecycle

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

A similar test as that which applies as for Freedom of Information requests should exist, complete with a list of exemptions.

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?

Similar to the FOIA local authorities should have the power to charge for information dependant on an assessment of the time that releasing the dataset would take. However, the value in publishing data which is frequently sought will never outweigh the costs of publishing it.

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?

A broad definition should apply – to those who collect and hold. Where information is held in numerous locations, (e.g. by both central and local government) the onus should be on central government to publish it, because of efficiencies in managing, handling, and making data available.

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

An appropriate mechanism might be linked to a presumption in favour of publication. In cases where datasets are not already published, they can be requested and then must be published subsequently – similar to ONS' "Commissioned Tables" approach. This could be enforced by the Information Commissioner.

A standard framework for publishing datasets in Local Authority organisations could be scoped by Central Govt and the ICO.

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

As above.

7. 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?

Whilst the ICO has enhanced powers, there should be a framework to ensure that the organisation has adequate policies and procedures to ensure that there are effective protocols and staff competencies in place.

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

Whilst the principles are in place, an amendment to the Data Protection Act should be considered due to the scope of the agenda.

9. 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?

There may be significant resource implications – related to developing or buying skills to improve online publication, software, etc. FOI tests on proportionality should stand. This is especially in the context of the need for tighter prioritisation because of shrinking budgets.

Unfortunately many Local Authority datasets were not designed to be distributed openly and were only meant for internal use. It will cost significant sums to clean the data up and publish it in a format that is of use to the public.

Whilst intending to increase the agenda on transparency, it is likely that increased data will highlight rights to information already available under FOI and DPA. It is likely that these types of request may increase, impacting resources further.

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

This may be dependent on any amendments made to legislation that can help to further ensure that contractors are data processors on behalf of the data controller.

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

Central Government should take the lead on software and publication. Strict guidance and use of meta-data should be in place across all publishing bodies. Local authorities should also ensure that close networks are developed to ensure localised standards, particularly in urban areas where boundaries are close.

Councils need a fixed standard for publishing data so consumers can compare like with like. For example, it may sound like a good idea to tell Council's to publish details of all their contracts, but that is of little value unless you know what the contracts cover and the public are able to compare that with the contracts of other councils.

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

More information would be needed on this proposal, including definition of 'user experience' and how this might work in practice.

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

More market testing would be required to see the demand from intermediaries in order to benchmark this.

14. 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?

Simplify requirements for returns and publications. Before implementing changes the regulations/ requirements on publication schemes, asset registers, and the single data list to name a few should be merged.

15. 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?

Each organisation should ensure that there is a senior information asset owner in place, and the same person should be responsible for ensuring compliance under existing legislation.

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

This should be led on by the ICO and central government, but should be proportionate to existing sanctions under DPA.

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

All sectors would benefit from having a dedicated transparency board - so this should be mandatory.

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

A single inventory for local government should replace and simplify the multitude of existing requirements for returns from various central government departments and bodies. This should be widely consulted on and be a cross-departmental project, involving all departments who have an interest in or co-own data held by local authorities.

Data is held by hundreds of Council's and Government departments, often in different formats. This makes it difficult for the customer to know where to go. If this data is to be of value – it would be better to be available from one standard portal where it can be collated together.

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

Data should be prioritised according to its usefulness in governmental decision making and its importance to achieving the goal of transparency.

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

Data about public services, user satisfaction and the performance of all providers should be collected and published - particularly aggregated, anonymised data by customer/citizen characteristics, in line with equalities considerations.

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

Certain data might be collected unnecessarily but more common is that public data is published and republished in numerous locations, which can make access confusing. An example might be benefit claimant data published by www.ons.gov.uk, and by www.nomisweb.co.uk, and on <http://neighbourhood.statistics.gov.uk>, and on www.dwp.gov.uk and from www.data.gov.uk – in this example it is not that collection should be stopped, but that access to the source with the most detail available should be streamlined.

There clearly are data sets which have a high value interest, of great policy or administration interest, or of information value, but does most of the data we hold fall into that category? If we publish too much data not only will the cost be exorbitant, but there could be information overload with the public unable to find the useful data among all the irrelevant data.

The requirements for collecting data could in addition be regulated through fields in inventories.

22. 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?

Quality should be linked to the useability of the data to make policy decisions. If the data is not of sufficient quality on which to base decision making, or be considered as a valueable source, it should not be released as it is not of sufficient quality.

If Councils undertake this project 'on the cheap' they will end up publishing raw data. If this happens it will only benefit large corporations who have the skills or the money to obtain the skills to interpret such data, and not the general public.

We hold so much data that we could overlook data that is of value to our customers. The data that is important to us may not necessarily be the same data that the public want to know. Therefore, some type of consultation on which data to publish would be required.

23. 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?

This should be held on a central portal with easy upload, ensuring all departments use this.

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

Their status as key data about public services, user satisfaction and the performance of all providers. Their usefulness in policy analysis and decision making, and ability to contribute to an understanding of public policy and social issues.

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

Central government should publish existing data at a more detailed level – e.g. cut in more ways, by equalities strands, at lower geographies, etc.

26. 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?

Specific competitions are unlikely to yield the kind of results to deliver the UK government status as a world leader in transparency. Improving accessibility, both in terms of increasing the use of machine readable formats and having strong and consistent meta-data standards, on a single government site will stimulate the use and reuse of data in and of itself. If work is done to improve this site, this should be strategic work by the UK government with world market leaders and innovators in web development, library science and open access data.

27. Overarching Comments:

The split between information and data is very important. Data is not necessarily accessible for members of the public, regardless of it's format, because it is not interpreted. The main audience, realistically for raw sources are businesses looking to sell analysis services to the public sector, think tanks, academics, other government bodies and other researchers. While this greater use of raw data may be positive, information also has a link to transparency, and transparency is about more than just publishing data. It is about honest evaluation, benchmarking, and analysis that adds value and understanding about a policy or service's relative successes and failings.

Finally, it should be recognised that there are limited resources in the public sector at present, and that if additional resources are required to be diverted to ensure greater access to data, then this risks impacting on the resources available to deliver frontline services. Furthermore, if resources are diverted into making available large amounts of new data, setting up a data corporation, etc, then there is a risk that there is less time and energy overall diverted to analyse and use it - as the professionals who analyse and use the data are often the same ones that collate and publish it.

The problem with publishing this data online is that large numbers of our residents, (particularly those who are most deprived) have no access to the internet. We need to tackle the issue of social inclusion in the digital age, to make sure all our residents are benefit recipients of this and not just the affluent ones with the necessary computer skills.

Ultimately, a balance needs to be struck between transparency and privacy. In many instances, Local authority datasets relate much more closely to the activities of the individual in comparison to many existing central government databases, and I believe that standards of best practice should be developed sector by sector, including collection practices.