

RSPB Response to the Open Data Consultation

Please find below my response on the Open Data Consultation on behalf of the RSPB. I greatly welcome this consultation and the Open Data approach. I am hopeful that it will be hugely beneficial for the UK across every sector, but I feel that it will particularly help the Third Sector. I have today also submitted a response to the consultation on Data Policy for a Public Data Corporation.

I have set out my response here in the order of the sections and questions in the consultation document:

Glossary of key terms

1. Do the definitions of key terms go far enough or too far?
 - I would extend the definition of **Dataset** to say not just “Factual data, structured or unstructured” but to say “Factual data, structured or unstructured, and related to location with a location identifier where possible”
 - I would extend the definition of **Open Data** to include data which may only be open for charitable, academic or other non-commercial use.
 - I would have a further key term in addition to the definition of **Public Services** which would be **Public Good**. This would be defined as, for example, public good provided by registered charities. Then Open Data can be aimed at enabling greater public good as well as better public service.
2. Where a decision is being taken about whether to make a dataset open, what tests should be applied?
 - Whether public release of the data would compromise **national security** or **personal privacy**, or cause **environmental harm** (eg criminal or inadvertant disturbance to rare species)
 - Whether the data would lead to an **improvement in the delivery of public services or public good**
 - Whether the data would **enhance the operational performance** of a sector of the UK: government sector, academic sector, 3rd sector, public sector, commercial sector, military sector etc
 - Whether the data are necessary for the Government to fulfil its **statutory duty**
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?
 - If the costs to publish arise from the publisher not yet being ready to publish, the costs should be borne by the publisher – to stimulate the publisher to implement a cost effective way to release the data they could reasonably be expected to be releasing.
 - The requestor should only be required to pay for public data if they require the data to be supplied in a bespoke manner, ie not through the common infrastructure implemented for INSPIRE compliance, or if a printed copy is required rather than a digital copy
 - If the costs are very high, the requestor should be required to pay if they need the data more urgently than the data can otherwise be available; my assumption is that ‘eventually’ all public data should be available – but this has a very long time line.

4. How do we get the right balance in relation to the range of organisations (providers or 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?
 - The definition of public services and public tasks should inform what range of organisations should be in scope. If it is a public service, the provider should be in scope; if it is a public task, the deliverer of that public task should be in scope. The key test should be whether the Public are, or could be, consumers of the service itself, and whether better public services or greater public good (without compromise to national security or personal privacy or the environment) could arise from the information being open.
 - A key criterion therefore is whether including the organisation as being in scope could in any way enhance the provision of public services or public good.
 - Whether the information arising from or relating to the public service can be considered to be a public asset – it would be if these information are key to the protection and operation of the UK, and health and wealth of its population, economy and environment, and to the effective function of Government and Public Bodies.
 - Data arising from organisations delivering public good or any part of the public task should also be in scope if not burdensome or costly to make available – it is a more delicate balance at this moment in time to widen the scope to include non-government organisations, but it should not be ruled out.
5. What would be appropriate mechanisms to encourage or ensure publication of data by public service providers?
 - The Inspire regulations and other applicable legislation
 - Formation of the Public Data Corporation so that key data creators are operated together and united for increased efficiency and effectiveness and greater public benefit
 - Removal of the principle of key Public Bodies operating as trading funds – this inhibits data from being open
 - Costs of servicing data requests is borne by the publisher so that it is more cost effective to publish efficiently and have data easily available
 - A change management programme – open data is a huge change and so there will be resistance and that needs to be spotted and targeted for attention

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?

6. How would we establish a stronger presumption in favour of publication that that which currently exists?
 - Start demonstrating the public value from having published, and accredit those who publish successfully and promptly
 - Start demonstrating the corporate value from having published and use intangible benefits as well as monetary benefits to calculate value and ROI
 - Discuss with publishers what would directly help them, and listen to what key enablers would be
 - Work with the Information Commissioner to address any failure to publish

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?
 - Yes- the Information Commissioner should not be duplicated, there should be a single body safeguarding the rights linked to data so that it is easy for everyone to know where to go
8. Are existing safeguards to protect personal data and privacy measures adequate to regulate the open data agenda?
 - I think so, nothing is changed as a result of more data being open, it should be business as usual but over a greater volume of data
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 its aim?
 - It could be quite large and burdensome during the early stages of the transition to this being normal working practice, particularly if ever the scope is widened to include academic bodies or the third sector; however once it becomes normal working practice, it should not feel quite so burdensome, and if ample time is allowed where needed to bring in older data that may take longer to compile into modern-day format, that will help things to feel feasible
 - If the PDC does open up all data it will lose immediate sources of income before greater uses lead to greater economic returns through taxation etc, or through income from any commercial uses of the data. For the PDC, this burden is proportionate to its vision of enabling public access to public data, and it is this body which must go all the way. For other public bodies, the transition could be phased, and the focus should be on priority datasets that will lead to maximum public use and value, rather than every possible dataset regardless of priority
10. How will we ensure that Open Data standards are embedded in new ICT contracts?
 - Simply specify the standards that will be required for payment to be released
 - Work with ICT suppliers so that they are familiar with the Open Data standards and supportive of them

Setting transparency standards: what would standards that enforce this right to data among public service providers look like?

11. What is the best way to achieve compliance on high and common standards to allow usability and interoperability?
 - Creation of tools, especially common or centralised tools, that only accept data that confirm to common standards
 - Checking of submitted data and recognition of their quality through 'kite marking' or similar
 - Promotion of high standards for public bodies who are going to receive greater scrutiny through the move towards transparency
 - Education of staff and users of the benefits of standards
12. Is there a role for government to establish consistent standards for collecting user experience across public services?
 - Probably – but I would initially only invest in creation of standards for a baseline

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

- Definitely – but it needs to be simple to implement and simple to spot so that users recognise and value the accreditation
- If the scheme is to work, it needs to be compulsory and across all sectors so that anyone involved in delivering public good or public services can be accredited, not just public bodies

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?

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?

- Use regulation to ensure that public service providers are clear in their legal duty
- Prosecute those providers who continue to shirk their commitments despite warnings and offers of support, so that after a given period of time to make the transition, transgressions are not tolerated
- Recognise the greater public good and public service improvements that the provider has enabled so that the provider is awarded a top rating that may able them to better compete for funding

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?

- It could include third sector representation to check as to whether data are penetrating the sector properly
- It could oversee change management activities to aid the radical change in culture expected (related to data being open); it could ensure that common standards are met across an organisation; it could ensure that the regulations are fully understood with a road map for all datasets held by the organisation to ensure they reach the public
- A different person should oversee the personal data safeguards in case there is a conflict of interest

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

- Yes if non-compliance becomes an issue

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

- All sectors would benefit, none should be exempt

Meaningful Open Data: how should we ensure collection and publication of the most useful data, through an approach that enables public service providers to understand the value of the data they hold and helps the public at large know what data is collected?

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

- Inventories are useful as baselines and there may be many uses of them in public services. This question is huge in scope – there may be an inventory of bus stops, an inventory of habitats, an inventory of who needs a particular service. I would relate it to Government objectives, public task and delivery of public good, not sure how else you would begin. The Annexes for Inspire list core geographies many of which would be potential candidates for having an inventory of. Some inventories may not be compiled by anyone, so Government may look to fill key gaps.

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

- Value should come from the inventory being better placed to help improve delivery of public services and public good if particular data are included. Public demand, or demand from public bodies in the course of fulfilling their public task, should drive which data should be included. RSPB of course would wish to see better inventories of habitats, wildlife and the marine environment. The marine environment is particularly data deficient and value would come from addressing this on behalf of all stakeholders with an interest.

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

- Across its service and policy portfolio, eg: health, crime, biodiversity, forestry, agriculture, education, equality, population, immigration, trade, climate change, environment, culture etc
- Across its spend and costs for transparency
- Across key performance indicators such as poverty indexes, biodiversity indexes, health indexes etc
- On the data it collects so that we can tell whether data are Open Data and whether everything that could be published is being published

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

- I don't know the extent to which data are collected that fall outside the public task but that which bring in income from commercial sales – arguably, these data may be unnecessary, but their collection has value; the question is whether Government should operate the collection, or the commercial sector
- If data collection is linked to fulfilment of government objectives and delivery of public services and public good, the data are necessary in this fulfilment and so should continue to be collected
- If the data are not publicly available (except where personal, of national security value, or preventing environmental harm), then arguably they should not be collected and collection should stop

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?

- The quality of data should always be stated, but quality can vary in lots of ways, so high quality is not necessarily consistent between different types of user. The real issue is whether the data contain errors – to the best of anyone's ability, data should always be error free and this is a key measure of quality, and whether the data have high quality metadata. As long as public data are relevant to the dataset they are in, and accurate and complete, and their metadata explain their currency and any attributes, then the data may be considered to be of high quality.

These are my measures of data quality:

1. Known provenance and provider
2. Availability
3. Currency
4. Completeness
5. Accuracy
6. documentedness (ie presence of high quality metadata)
7. format and its suitability for use and interoperability
8. detailedness of spatial resolution
9. presence of a unique identifier on each record
10. decipherable codes and attributes
11. meeting of common standards such as standard file formats, uk Gemini 2 etc
12. level of summarisation

- So data should always be of high quality in terms of accuracy and completeness and the quality measures above, but not necessarily high quality in terms of level of spatial resolution.
- By data polishing, I take this to mean ensuring that the data are complete and not misleading, with obvious anomalies spotted and fixed. It should not mean analysing and summarising – raw data should be just as available as summary data – and often data suppliers are worried about people drawing the wrong conclusion from data and wish to interpret data and just to provide interpreted data. Uninterpreted data are needed too, so that people can rework analyses from first principles to check for errors or to try alternative methods or purposes. But data should always be cleaned for errors, omissions and undefined codes. So data providers should polish the data to this extent.

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

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?

- A central portal would seem the ultimate aim, with departmental portals only necessary as a transitory step towards centralisation. Centralisation should bring economies of scale, version control and common standards.

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

- There should be a balance between datasets that bring *greatest* opportunity for improvement of public services or delivery of public good, and datasets that bring *earliest* opportunity for improvement of public services or delivery of public good
- In addition, public interest should also be served, so if in a local area there is concern over health issues, health data should have higher priority than say crime data

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

- Government should first aim to publish core reference data, especially spatially explicit data, that underpin interoperability, and then to publish a broad set of data, and then to publish more detail in higher priority areas.
- The key is to identify where the greatest public good and improvement in public services can be realised, and where data enable Government to fulfil key public tasks. These should be the highest priority data, and I do not know whether this would identify a broad range of data or some very focused areas of data.
- Because part of the Open Data agenda is to further innovation, a broad set of data may allow innovation to begin in a wide range of areas. Whichever way it is done, the end goal of all public data being public will be realised eventually.

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?

- To make all public data publicly available so that they are openly available for any purpose for free, including all PDC data. This would bring opportunities to every sector, to every organisation and to every individual. Competition would naturally occur between those exploiting the data and bringing products to market; new niches would spring up replacing some of the opportunities that will no longer be profitable because of the new openness for all public data. The lifting of restrictions on use of the data and of charges for the data would be hugely enabling and would allow absolutely anyone to innovate. It is the cost and licencing regimes of data at present that restrict innovative use. You want to say “just give us the data!” – because there are bound to be opportunities that you don’t even think of til you have the data to start playing with and exploring with.
- To promote use of the data at all times so that use is absolutely maximised, supported, informed and widespread.
- To expect the PDC to have included as part of its public task the role of stimulating innovation through data being as open as possible for both commercial and non-commercial uses
- To stop being worried that third sector use of data for free would adversely compete with commercial activities. This completely prohibits access to data for the third sector, and the third sector is only interested in delivering public good and undertaking activities that aid them in achieving this. Start off being open and regulate the use of the data later if it becomes apparent that charity use of data is a problem. The advantages of the third sector having full and open and free access to data will hugely stimulate that sector to be innovate; to date, it has been prevented from being so by licencing regimes that exclude and prohibit. We are up for being

UNCLASSIFIED

innovative and just need the Open Data world to enable us to have the open access to data we are dreaming of.

I would be very happy to provide more feedback on the Open Data consultation on behalf of either RSPB, the Third Sector, or myself as a private individual. Please let me know how I may be able to help.

I look forward to hearing about the outcomes of the consultation in due course.

Best wishes

Ellen

Ellen Wilson

UNCLASSIFIED