

RCUK RESPONSE TO THE CABINET OFFICE CONSULTATION ON MAKING OPEN DATA REAL

1. Research Councils UK (RCUK) is a strategic partnership set up to champion research supported by the seven UK Research Councils. RCUK was established in 2002 to enable the Councils to work together more effectively to enhance the overall impact and effectiveness of their research, training and innovation activities, contributing to the delivery of the Government's objectives for science and innovation. Further details are available at www.rcuk.ac.uk
2. This evidence is submitted by RCUK and represents its independent views. It does not include, or necessarily reflect the views of the Knowledge and Innovation Group in the Department for Business, Innovation and Skills (BIS). The submission is made on behalf of the following Councils:
 - Arts and Humanities Research Council (AHRC)
 - Biotechnology and Biological Sciences Research Council (BBSRC)
 - Engineering and Physical Sciences Research Council (EPSRC)
 - Economic and Social Research Council (ESRC)
 - Medical Research Council (MRC)
 - Natural Environment Research Council (NERC)
 - Science and Technology Facilities Council (STFC)
3. This response focuses on data generated by the Research Councils in the course of their activities in supporting research and enhancing access to research outputs and infrastructure. Therefore it does not address every question in the consultation.
4. The Research Councils are committed to transparency and already publish information about the research they fund, research spending and overall success rates. Publishing this information, as well as providing more detailed information directly to Research Organisations plays an important part in informing decision making, for researchers and research organisations, on submission.
5. RCUK's position statements on access to research outputs and the common principles on data policy¹ also outline the Research Councils commitment to enhancing access to outputs of the research they fund through ensuring publications are openly accessible, and enhancing the value of research data either through publication or data-sharing between researchers to support new research. To support this Research Councils fund infrastructure which help to delivery policies such as the UK Data Archive and UK PubMed Central (UKPMC).

Glossary of key terms

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

6. The definitions are broadly appropriate and provide a good starting point for generating public debate. However, we note that there is no definition provided for "Public Data". It will be important to clarify what this covers since it is not clear that the working definition given in Annex 2² allows scope for different interpretations of what is expected of data produced in the course of publicly funded research.

¹ <http://www.rcuk.ac.uk/research/Pages/outputs.aspx>

² Annex 2 - objective, factual, non-personal data on which public services run and are assessed, and on which policy decisions are based, or which is collected or generated in the course of public service delivery.

7. Likewise it will be important to define “personal data” and ensure that this definition covers data that can be unambiguously related to an individual or an organisation. Anonymised personal data plays a crucial role in research especially in social science and biomedical fields. Personal data is also used (subject to the Data Protection Act) in research administration (grants, etc).

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

8. A widespread benefit for the public should be the main testing point for making data open. There is a wealth of government-collected data which could form part of the infrastructure for research e.g. social science research, which informs policy. As part of the decision process careful consideration must be given to whether or not data should be made open. Where this is not possible, the value of making the dataset discoverable and providing information on any restrictions that may apply should also be considered.
9. It is not enough to make data open simply because they are available; open data must be of appropriate quality and represent value to the public. Any data that are released should have a quality statement from the producers and relevant contextual information to prevent or reduce the possibility of misinterpretation. This is critical to maintaining public trust in the integrity of data which informs policy decisions.
10. The exemptions outlined under Freedom of Information legislation provide useful tests for openness as they call for the consideration of issues such as information intended for future publication (which would usually include quality and contextual information), commercial interests (including intellectual property considerations and personal information (reflecting the requirements of the Data Protection Act). FOI can also act as a useful tool to opening up further datasets not initially identified by organisations as being candidates for Open Data. RCUK considers that once a dataset has been released under FOI, it should be considered as Open Data and that updates should be published routinely as for other Open Data.
11. It is important that openness (data sharing) is pursued not as an end in itself, but to maximise the value of the data and the ultimate benefits to the public. This requires custodians of data, and those who wish to have access, to understand the data lifecycle, when in that lifecycle sharing best adds value, and the risks associated with inappropriate access (e.g. to confidential information). For instance, premature sharing and analysis based on incomplete clinical trials data would not only be poor research (the trial may need to be stopped before it reached a clear conclusion), but may also have significant impact on the participation of patients and volunteers; it may also raise ethical concerns. The imperative to share data must not compromise research outcomes and quality, the legal and ethical requirements relating to the confidentiality and consent, or public confidence in the research.
12. Making data about society and government available is powerful for users of public services in very direct ways (suggested in the consultation paper under the *accountability* and *choice* opportunities), but the broader understanding which may bring about policy does not generally arise from just simple access to data. There are many situations in which complex data need to be subject to sophisticated analysis to make them understandable. There is a key role here for the research community, which is best characterised as research that contributes to *social* and *economic growth* through fostering public debate and informing policy decisions. For example, the police.gov.uk data represent a valuable asset, but they also raise considerable opportunities for misrepresentation and do not provide users with guidance as to what might be appropriate

denominators and reference areas for the statistics delivered. Indeed, some suitable information might be available elsewhere in the Open Data domain, but it should not be assumed that the release of raw data will lead to methodologically appropriate use or improved public understanding and more informed public choices. Hence, it is not just data release that matters, but the act of adding value to data that will make it possible to exploit the six opportunities, outlined in the consultation paper.

13. In the case of research data, it will be important for the procedures to be adopted not to require premature release of data prior to their appropriate cleaning to ensure high quality and sometimes to enable use in research publications or innovations, in order to avoid making research organisations give away competitive advantage in their research. However, as for any data produced by public services, there should be a clear and limited timeframe within which data must be released and justification in place if immediate release is not possible.
14. RCUK expects research organisations to take effective decisions about intellectual asset management to deliver the most benefit to society and the economy. This includes recognising circumstances where the publication of research outcomes or free dissemination to users might be the most effective approach. RCUK recognises the convention that the benefits should accrue to the individuals and organisations that create value, and RCUK's Knowledge Exchange Principles³ provide clarity for researchers and research organisations on the RCUK position on intellectual property and asset management.
15. In the case of Research Council administrative datasets, it will be important for the tests to consider the duty of confidence the Research Councils owe to applicants and research organisations and the responsibility the Research Councils have to preserve the confidentiality of the detailed information provided to support peer review which might include personal information, preliminary research plans, hypotheses and data, intellectual property etc. change

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?

16. In considering this issue it is useful to look at examples of how the research councils charge for data and information. The ESRC's Economic and Social Data Service (ESDS) provides free access to its data with no discrimination in terms of users (unless there are restrictions imposed by data licences or other access restrictions). However, the ESDS applies charges if the data are used for commercial purposes in accordance with its charging policy. The NERC policy is to make the environmental data it holds available free to all users, apart from large or complex requests or where a request involves third party data where NERC itself is required to pay a fee. NERC does, however, charge on a cost recovery basis for a number of value added information products, such as ground stability reports or flood risk maps, and is currently finalising a new policy on how and when it will charge for information products. NERC's aim is to supply information products for free or at greatly reduced cost to support non-commercial teaching and research, private use, and innovation activities. All of NERC's charges are compliant with current HM Treasury guidance. In addition, where appropriate NERC's research centres either have, or will be working towards, Information Fair Trader Accreditation from the Office of Public Sector Information.

³ <http://www.rcuk.ac.uk/kei/expectation/Pages/kePrinciples.aspx>.

17. RCUK considers that the NERC model of mainly supplying data free of charge, but charging on a cost-recovery basis for value added information products should be adopted by all public service providers. If charges are made for data they should be on the basis of recovering the costs of meeting the request for data only.

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?

18. A key consideration here should be the potential demand for the data from various users (commercial organisations which see the scope for value-added products, the research community which identifies their research value and members of the public who have an interest in the activities of public sector organisations).

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

19. The most important driver is an overall agreement or consensus among the concerned parties on the significance of publicising data for public good. In research a culture of data sharing is well established and widely supported; this provides a solid foundation for any further mechanisms, such as suitable infrastructure and common data and metadata standards.
20. Individual Research Councils have policies and support infrastructure both to promote research data sharing and to ensure proper management of intellectual property, including IP derived from research data and software. Most Research Councils also require applicants to submit data-management and data-sharing plans as part of their funding proposals, as do The Wellcome Trust, Cancer Research UK and other research charities. Several of these organisations are currently reviewing the requirements of these plans and how they are used in peer review, to ensure they are fully understood by the research community and effectively promote quality data management and sharing. Similar requirements might be used by other public bodies.
21. All the Research Councils already make details of awarded grants publicly available by clarifying their expectations around the level of detail that would be routinely published at the time of application. The Research Councils also actively review how they might improve the dissemination of information about the portfolio of research that is supported as well as the outcomes of that research.

An Enhanced Right to Data

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

22. By recognising the contribution of an individual or team in producing the publication/dataset product via reward systems such as for example promotion, tenure and esteem use as a criterion in assessing research proposals etc. There is a particular need to ensure that those researchers who invest in generating new data resources get appropriate credit for active and early data sharing, especially in the Research Excellence Framework (REF).

23. RCUK's Common Principles on Data⁴: individual Research Council data sharing and management policies recognise that data arising from Research Council funding are a public good and should be made openly available with as few restrictions as possible in a timely and responsible manner. RCUK recognises that there may be legal, ethical and commercial constraints (including issues relating to intellectual property) on release of research data that does not harm intellectual property. We would expect researchers in receipt of Research Council funding to comply with these policies, and the associated policies on open access and deposition of research publications in public repositories (e.g. UKPMC) where appropriate.
24. Individual Research Councils have policies which elaborate on the common RCUK principles for their fields. For example, for large scale facility based science, where data are collected and stored centrally, it is possible to automate the release of data after an embargo period if appropriate. The STFC ISIS facility has implemented such a policy.
25. RCUK believes that the existing legislative framework not only already establishes sufficient presumption in favour of publication but is overly complex (please see response to Question 3 below).

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?

26. Whilst recognising that the FOI legislation is a foundation for openness and transparency in government, the Research Councils are concerned that the FOI legislation is often perceived as giving insufficient recognition to the special nature of research data produced in the course of normal scientific investigations. The Research Councils would therefore welcome enhanced powers allowing the ICO scope to ensure that research data are not prematurely released to the detriment of the research process whilst also having powers to ensure that such release, where appropriate, is not unnecessarily delayed. There are likely to be many marginal cases requiring sensitive adjudication and it is particularly important to note that discipline specific expertise may be required to reach appropriate decisions in some cases.
27. Furthermore, it is important that any independent body with powers and scope to safeguard rights to data and data access understands both the opportunities and potential threats to openness and confidentiality that the new technology introduces. Such understanding may be essential in any consideration of how the law will need to be applied or modified.

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

28. The legal framework in the UK in relation to personal data and sensitive personal data is complex due to the overlapping but at times divergent requirements of statutes such as the Data Protection Act 1998, NHS Act 2006 and the common law precedents on consent and confidentiality. This can make disclosure of data, even if appropriately anonymised, difficult and can be a disincentive – real or perceived – to data sharing. The complexities of this situation (and recommended solutions supported by RCUK) are described in the 2008 Data Sharing Review report by Walport and Thomas⁵.

⁴ <http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx>

⁵ <http://www.justice.gov.uk/reviews/datasharing-intro.htm>

29. Medical researchers in particular have been concerned that the UK legal framework for data protection and confidentiality is ambiguous. As a result, the law is sometimes unnecessarily interpreted in ways that restricts sharing and consequently the potential to benefits that arise from data sharing. It can be argued that the protection in the Data Protection Act given to “personal information” (as distinct from “non-personal”) is unhelpful, and that the better focus would be on protecting information that is entrusted “in confidence.”⁶ There are shared concerns in the social science and medical communities that access to and linkage with government and NHS administrative data are unduly hampered in these ways, although such linkage is of demonstrable benefit to social, economic and health research and thus benefits society. Safeguards in access and usage of such information need to be appropriate to ensure confidentiality, but not overly restrictive.
30. RCUK acknowledges the complexity of the regulatory and legal framework governing research data in the UK, which includes significantly different treatment of research data in Scotland, is a cause of concern and anxiety to many publicly-funded researchers and private sector research users who collaborate with them. In their report on the handling of climate research data at the University of East Anglia, the House of Commons Science and Technology Committee concern expressed the view that “the broader confusion about how FoI legislation should be applied to scientific research must be resolved”. RCUK endorses this view and welcomes Andrew Miller MP’s draft amendment to the FoI which related to research data intended for future publication currently proposed under the Protection of Freedoms Bill. RCUK also welcomes the ICO’s new guidance on the disclosure of research information.
31. Poor data security impacts negatively on public perception and while a robust legal framework is crucial, the importance of physical and computational safeguards and appropriate institutional policies should not be underestimated in assuring the protection of personal data.

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?

32. The costs (both in terms of financial cost and time cost) associated with setting up and maintaining large data sets can be high. A rule of thumb estimate used in some universities to cost the provision of storage and backup of research data is approximately £1/gigabyte/5yrs total cost, excluding extended data curation. While data volumes vary enormously across disciplines and institutions, a rough approximate calculation could be that an institution with 2 to 3 petabytes⁷ of research data would need to set aside approximately £2m to £3m over a 5 year period. However the costs of storage and back-up account for only 5 - 10% of the cost of preservation, depending on the discipline. The second ‘*Keeping Research Data Safe*’ study conducted by Neil Beagrie et al, 2010 and funded by JISC identified that approximately 40% of the cost of digital preservation⁸ is spent on adding information about the data (metadata) before it is archived and the remaining costs are incurred in mechanisms to find and access the data, particularly for non-text data. Furthermore, a credible private sector estimate obtained by the research councils (that it costs around £22 million to preserve 1 petabyte for 5 years) suggests that the above may underestimate the true cost.

⁶ ‘Can better use be made of public data for example in health research?’, Foundation for Science & Technology, 8th June 2011

⁷ kilobyte (kB) 10³; megabyte (MB) 10⁶; gigabyte (GB) 10⁹; terabyte (TB) 10¹²; petabyte (PB) 10¹⁵

⁸ <http://www.jisc.ac.uk/media/documents/publications/reports/2010/keepingresearchdatasafe2.pdf>

33. RCUK believes that it is legitimate to use public funds to support the costs of storing and sharing publicly funded data, and that the costs of preparing data for storage and depositing it in an appropriate repository should ideally be covered as part of the project costs. However, these costs are increasing significantly, not only due to the growing volume and complexity of research data, but also due to the need to make data and metadata more widely accessible to lay audiences, policy makers or researchers in different disciplines. In a period of financial constraint there is a risk that research programmes may have to be reduced to allow more resource to be directed towards data management (though this risk is to some extent balanced by the potential research efficiencies that come from sharing rather than creating fresh research data). There is also the risk that as data become more readily accessible; short-termism might reduce the incentive to fund the collection and maintenance of very high-cost data even though this may hold great potential value.

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

34. We welcome the principle that data should be open by default. New ICT contracts covering the digital storage of open data should clearly define appropriate metadata standards to ensure that the public may easily discover the existence of such data (and any applicable conditions of access) using online search engines.

Setting Open Data standards

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

35. A balance needs to be struck between the breadth of disciplines for which a standard is applied and the time taken to establish that standard into practice. Where the data involved are similar, harmonisation of approaches to data collection by public bodies will do much to promote subsequent interoperability and opportunities to derive new knowledge by combining existing data.
36. An approach of awarding contracts / grants only to those proposals which state they will comply with the stipulated standards and by barring those that do not subsequently comply from bidding for further contracts/grants for a given period would help to achieve compliance.
37. For research data, rigorous peer review of data management plans can play an important role. RCUK has concerns that where standards are developed and implemented through the International Standards Organisation (ISO) they are made exclusively available in the UK through the BSI, and are only accessible through payment of an appropriate licence fee. Indeed, Research Councils invest time and resources in contributing to the development of international standards, only to then have to pay a second time to access those standards via the BSI. This 'commercialisation' of standards is a barrier to uptake and use of appropriate standards, especially within the research community.

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

38. The Research Councils recognise the need for consistent standards in public data to enable secondary users to fully utilise such information.

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

39. As information intermediaries become a significant force in interpreting “open government data” for clients, there will indeed need to be regulation and so accreditation for bodies which hold data on behalf of others. For scientific data, there are existing schemes for the certification of trusted digital repositories at National, European and Global levels⁹.

Corporate and personal responsibility

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?

40. RCUK fully supports Open Data. Individual Research Councils support domain specific data repositories that provide access to their data holdings. Although in some cases there are access restrictions, for example due to licensing rights or privacy arrangements, but in general, these data repositories also provide access to non-academic users i.e. to users in the third sector or the commercial sector as well the general public. For example, the ESDS funded by ESRC provides generally free access to over 5,000 digital data collections, including data generated by public service organisations.

41. RCUK recognises that with more detailed and secure data comes more responsibility for their safety whilst maintaining accessibility. For social science data the ESRC funds the Secure Data Service – a unique service in the UK with its philosophy of enabling secure and safe and yet remote access to sensitive data produced not just by the data collection it funds via research, but also by public and private sector organisations.

42. There are legal and ethical constraints on release of personal information gathered as part of research. Public trust in researchers and confidence in the research process is absolutely crucial. To that end, the Research Councils promote high levels of accountability, transparency and good practice in research involving personal information,¹⁰ including compliance with international standards and government guidance for information security practices. To ensure that the research process is not damaged by inappropriate release of data, research organisation policies and practices should ensure that these are considered at all stages in the research process. A key aspect of this will be controlled access of researchers to different levels of raw data and the anonymisation of such data before wider release (see above for information on the Secure Data Service).

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 are properly protected and that privacy issues are met?

43. No comment from RCUK on this issue.

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

44. No comment from RCUK on this issue.

⁹ <http://www.dcc.ac.uk/resources/curation-reference-manual/chapters-production/audit-and-certification>, <http://www.trusteddigitalrepository.eu/Site/Welcome.htm> and <http://public.ccsds.org/sites/cwe/rids/Lists/CCSDS%206520R1/Attachments/652x0r1.pdf>

¹⁰ For example <http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Dataaccess/index.htm> and <http://www.esrc.ac.uk/about-esrc/information/data-policy.aspx>

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

45. No comment from RCUK on this issue.

Meaningful Open Data

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

46. Data inventories are a useful tool for acknowledging the broad collection of data created by public services in a coherent way. Being created in interoperable format and coordinated they would provide very useful information on gaps and overlaps in data collected by public services leading to potential efficiencies. Data inventories may include information on data use and purpose of such use to inform public services collecting data on their actual use and hence the rationale for further collection (e.g. if use of certain datasets is very low or non-existent, further data collection may not represent a good value for money).

47. It is important that the appropriate metadata are available so that data used for the research itself are presented in context. This is particularly important if the data are to be used by others that are not expert in that particular field including researchers from other fields.

48. Data inventories should also include enough "contextual information" for any user, from any domain, to be able to understand how the data were collected, how the data have been manipulated before analysis, what analysis has been carried out already (i.e. links to publications) and an indication of the quality of these data. The provision of this information will also reduce the possibility of the data being misinterpreted.

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

49. There is a strong argument – based on the concept of research being built on solid foundations of previous research – that research outputs (publications and data) should be available openly and free at the point of access. The costs of replicating research (other than for confirmation or elaboration of earlier results) usually are higher than re-using/re-interpreting existing information. Thus access to open and freely accessible research resources can promote quality and can be both effective and efficient – but only if the metadata are sufficiently rich to expedite discovery and to facilitate further utilisation.

50. User feedback and the level and purpose of usage could be a useful tool to inform data prioritisation in an inventory. Therefore establishing a mechanism whereby the user can provide feedback to the data provides is an important element of Open Data.

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

51. There are many areas of government where certain data should be collected and published routinely. For example, in the social and economic sciences we consider that any data which have the potential to 1) drive innovation or progress social and economic growth; 2) reduce inequalities in all areas of life; 3) combat poverty; 4) promote individual and nation's health and well-being 5) understand the causes of conflicts and their possible prevention mechanisms and 6) explain how markets work, should be prioritised and be made available for informing public policies.

52. The areas should be also informed by user needs for data and help inform their choices. Hence, mechanisms for user feedback (mentioned above) are crucial.

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

53. No comment from RCUK on this issue.

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?

54. Public data should be released in a suitable and standardised format to allow easy and useful discovery, re-use and re-purposing and linking with other data. There is little use in releasing data of such low-quality that it cannot be taken as statistically significant for the uses to which it may be put. However, where 'polishing' takes place prior to publication this should be clearly indicated when the data are made available so that those using the data are fully aware of its provenance. It is also important that, whatever the quality of the data, the relevant metadata, including machine-readable metadata needed for automated semantic mining, should also be published to ensure the resources can be used in context.
55. As mentioned above in regard to decision making on whether to make a dataset public, there is an element of trust in data produced by public services that could be undermined if low quality data were released. If such data were used in research it would undermine their integrity and also ability to inform policy and drive innovation.

Government sets the example

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?

56. In certain areas (e.g. social science, medical) there are already clear examples of social policy being determined by the availability of open research outputs. Making public data (e.g. on health, education) available to social and medical researchers can be the starting point for new investigations, which can lead to improved public policy that can have wider benefits for society. The Government's initiative on making public data public is a major step to provide a platform for free re-use of national and local data to facilitate the advancement of public services across the country.
57. There is a vast array of untapped data collected by Government departments which, if more widely accessible, could be used to improve the health and wealth of the nation. Linking such administrative data with ongoing surveys (e.g. birth cohorts, panel surveys) offers potential for unprecedented scientific and policy impact through tackling issues that, without such access and opportunities to link data, are much more costly and difficult to answer (for example: educational and occupational mobility; health progression and potentially understanding pathways to well-being).
58. Data linkage offers an additional and often highly powerful alternative to traditional data collection, whilst addressing concerns around reducing response rates to surveys via enabling calibration of data sources against each other. Enhanced data generated through linkages add to the evidence base and thus helps improve the quality of public policy and more efficient delivery of public services. Other benefits that can be realised through data linkage enabled by data access include the evaluation of public policy

interventions in ways that are not currently possible; for example, this would help us understand more about the localism agenda and the 'Big Society'. Furthermore, the use of linked administrative data has the potential to inform rolling population counts for small areas and so would lead to much more efficient methods for distributing the public money being spent on services in proportion to population numbers. The ONS is already exploring what alternative data might be used to provide such counts if there were no 2021 census, and the ESRC is supporting developments through a number of collaborative targeted initiatives.

59. In support of its strong commitment to data sharing and openness of data access, the ESRC is engaged in discussions on transparency of government data under Open Government License. ESDS has been providing access to government data, first and foremost produced by the ONS. ESDS is working closely with the ONS to enable easier data access with the simplified registration process to allow for more users (particularly non-academic) to have a seamless access to data.
60. Drawing on the above examples, we consider that a departmental level would be more appropriate, as they will have the knowledge on the data.

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

61. Examples from the research sector provide useful examples for open data. In the research domain, datasets identified for publication are prioritised by the volume of downloads of related publications and/or and requests for the data. Priorities are also linked to the actual or expected economic and research impact of research data.

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

62. See response to question 2 (above). It is probably more important to prioritise material which is not yet available but for which there is a demand.

Innovation with 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?

63. There are enormous opportunities to stimulate innovation, but they are dependent on the right Open Data being easily available to those users motivated to access and use it for innovative purposes. Simply having more data will not in itself produce an increase in innovation.
64. Government could consider using competitions and prizes to stimulate innovation in this area, but it needs to be tailored. For example, BBSRC's Innovator of the Year competition¹¹ celebrates the successes of BBSRC-supported scientists in delivering economic and/or social impact from their excellent research.

RCUK – October 2011

¹¹ <http://www.bbsrc.ac.uk/business/impact-incentive/innovator.aspx>