**Research Sector Transparency Board - Minutes**

**1 Victoria Street, Ministerial Conference Room B, 7th floor, 1030-1230 8 May 2014**

**Attendees:**

Members:

The Rt Hon David Willetts MP (Chair)

Professor Geoffrey Boulton - Royal Society

Mark Thorley – NERC (deputising for Rick Rylance, RCUK)

Dr David Sweeney – HEFCE

Professor Adam Tickell – University of Birmingham

Professor Sir Nigel Shadbolt – University of Southampton/Open Data Institute

Guy Goodwin – UK Statistics Authority (deputising for Jil Matheson)

Krisztina Katona – Cabinet Office (deputising for Olivia Burman)

Guests:

Roly Keating – British Library

Nicola Perrin – Wellcome Trust

Professor Nick Wright – Newcastle University (representative for Russell Group working group on open research data).

Observers:

Maja Maricevic – British Library

Tim Bradshaw – Russell Group

Secretariat:

Ron Egginton – BIS

Anna Macey – BIS

Apologies:

Sir Mark Walport – Government Chief Scientific Adviser

Iain Gray – Chief Executive, Technology Strategy Board

Professor Lynn Gladden – Pro-Vice Chancellor for Research, University of Cambridge

**Note of meeting:**

1. **Updates on open research data activities were given from the Russell Group, Wellcome Trust, RCUK, Royal Society and the British Library.**
   1. *Russell Group (RG) update given by Nick Wright, Newcastle University. Presentation available on the* [*RSTB gov.uk page*](https://www.gov.uk/government/groups/research-sector-transparency-board)*.*
      * Highlighted that the practicalities of open research data are a concern (in terms of the potential time demands on researchers and costs for implementation).
      * RG universities are great users of data - both within universities and from externals groups.
      * Foundation of open data begins with good data management.
      * RG encourages all RG Universities to be open and publish data management plans online.
      * Recognise importance of people. Need for training and making tools for data sharing available.
      * Importance of learning from best practice.
      * A concordat would encourage good behaviour.
      * Raised question of when should data be made available. There is a need to authenticate data. Working group to look at a decision tree for making data available.
      * Discoverability – need to make data discoverable otherwise limited use of open data practice.
      * Need to get realistic national cost.
      * Key Principles for open research data detailed in the RG presentation.
   2. *Wellcome Trust (WT) update given by Nicola Perrin. Presentation available on the* [*RSTB gov.uk page*](https://www.gov.uk/government/groups/research-sector-transparency-board)*.*
      * WT meet costs of data management plans included in grant applications. Data management plans need to be properly addressed.
      * Highlighted that there are distinct challenges between disciplines.
      * The Expert Advisory Group on Data Access has produced a report on Incentives and Culture Change for Data Access. The [report](http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/EAGDA/WTP056496.htm) was published on 30 May.
      * Clinical Trial Data – funded programme of work with Institute of Medicine and seeking to build an international consortia.
      * Agreed Concordat approach could be real step forward.
      * Sustainable infrastructure very important but need to consider costs.
      * Governance is key - but need to build public trust.
   3. *RCUK update given by Mark Thorley. Presentation available on the* [*RSTB gov.uk page*](https://www.gov.uk/government/groups/research-sector-transparency-board)*.*
      * RCUK open data activities are broken down into two parts: What the Research Councils are doing as funders and what the Research Council institutes are doing.
      * Sensitive to release of public data.
      * RCUK are producing guidance notes to tie in with RCUK common principles on data policy (<http://www.rcuk.ac.uk/research/datapolicy/>).
      * RCUK are trying to work towards a common policy on open data however this will reflect differences between disciplines (an example is the difference between the openness of environmental data compared to that in Social Science).
      * Highlighted importance to put effort in to metadata to make data useful.
   4. *Royal Society update given by Geoffrey Boulton, University of Edinburgh. Verbal update.*
      * Gave summary of Open Data Forum held on January 21-22 2014 (main meeting note available at: <https://royalsociety.org/~/media/Royal_Society_Content/z_events/2014/2014-01-21-Forum-Note.pdf>)
        + Priorities that arose from the Open Data Forum:
        + Need for a Concordat
        + Motivation – DataCite (for more information on DataCite see British Library update)
        + Need for Pilots.
        + Promote discussion of how data sharing is reviewed in assessment of grant applications.
        + Data ecology - need to express in clear language.
      * Meeting between representatives from the US and the UK:
   * Rational - science is an international enterprise. UK should not go it alone but should collaborate with others.
   * Aim of US/UK meeting was to ask if a parallel approach could be achieved. Could there be a shared set of key principles.
   * Identified a series of collaborative projects.
   * Dual approach across two countries. There is a need to think in broader international terms.
   1. *British Library (BL) update given by Roly Keating. Verbal update although a note covering two update topics is available on the* [*RSTB gov.uk page*](https://www.gov.uk/government/groups/research-sector-transparency-board)*.* 
      * Data Citation using DataCite:
        + Global infrastructure for which the BL is a founding partner.
        + Based on ensuring common citation identifier – DOI (Digital Object Identifier).
        + 3.2 million DOIs have been assigned globally.
        + British Library provide UK front end of DataCite.
        + BL would like to encourage increased uptake of DataCite use and develop its capacity so that it can support national uptake.
      * Webarchiving of Research outputs:
        + Aim to preserve material.
        + *Non print legal deposit legislation* enables five UK institutions and one in the Republic of Ireland to ingest everything on UK web but only available on the premises of these institutions.
        + *UK web archive* - permissions based - available on web.
        + Many UK HEIs and research institutions allow content to be archived on BL UK web archive, but the permissions based system is laborious to implement and might not fully reflect UK research strengths.
        + Systemic uptake of open web archiving would enable preservation of research content on the web, help with finding solution for ‘link rot’ and development of citation systems for web content.
2. **Following the updates the group discussed the topics raised during the first part of the meeting and also what should the next steps for Open Research Data (OD).**

* Issues/**Actions (in bold font)** which were raised:
  + Recognised that data could not all be in one place.
  + There is a need to recognise discipline differences and the need for different approaches.
  + Real need for pathfinder projects. These could explore how an open research data system might actually work - need to move fast to maintain national and international momentum. However there are questions over available funding. **Nick Wright to produce a paper outlining potential pathfinder projects in a number of sectors/disciplines and what the cost and merits for such a programme would be [ACTION].**
  + Suggested that budget for open research data pathfinders may need to be a mix of Capital and Resource funds, subject to what costs may be capitalised. BIS/ESRC could advise on how ESRC data projects are capitalised.
  + **David Willetts suggested that the group respond to Capital Consultation with an open research data bid for capital [ACTION].**
  + Skills provision is as crucial to pathfinders as capital funding.
  + Although open research data is another burden on researchers the group remarked that people are persuadable due to the potential gain for science. There are advantages to be gained by OD facilitating the validation of science and its improved commercialisation.
  + Even so, we must also remain mindful of the cost of OD.
  + Datacite was agreed to have the potential for wider take up, but in an elective and not a prescriptive way.
  + Emphasis should be on advocating the direction of travel to encourage wider adoption of Datacite.
  + Voluntary increased uptake of DataCite should be incentivised. Recognition should be given to researchers who use the service and Open Access data journals.
  + Geoscience Data Journal - Wiley or Scientific Data – Nature Publishing Group are illustrative of how scientific data can be formally published and the dataset creator attains full credit for their efforts.
  + Open Access data journals provide the community with full version control making major datasets citeable, discoverable, interpretable, reusable and reproducible.
  + The group considered that a Concordat will provide researchers with a structure to make decisions about data management.
  + **A Concordat developed in conjunction with The Russell Group and RCUK is likely to be published in the autumn [ACTION].**
  + The next meeting should focus on a discussion on potential pathfinder projects.