Royal Society Proposal for an Open Data Forum

It is proposed that the Society's final formal act in promoting implementation of the recommendations of its report "Science as an Open Enterprise" should be to convene an "Open Data Forum". The purpose of the event would be to engage with interested parties, whose attitudes and actions will be key to implementation, to identify how progress could best be made and how barriers removed or circumvented. If the participants wished to maintain a forum beyond the proposed event, it would be for them to do so.

Context
1. The RS Report "Science as an Open Enterprise" identified those parties that had key roles and responsibilities in realizing an effective regime of open data in research funded from public sources. These are those that fund science (RCs/FCs), those who employ scientists (universities/institutes), learned societies who influence the principles and priorities of scientific disciplines, those who publish the outputs of science, government, those that use science and not least, scientists themselves.

2. It has become clear that the report was timely, and has succeeded in re-energising the engaged community and stimulating other interested parties to action. It has also reinforced the government's policies for open data. The actions that have followed in the UK from the disparate groups listed above are:
   • Universities (supported by JISC) are actively seeking to develop their own policies and create a road map for open data.
   • Research Councils have strengthened their view about the importance of the issue.
   • Government has created a Research Sector Transparency Board to address how best to make research data open.
   • The British Library and University Libraries (through JISC and LIBER) are actively considering their role in a "post-Gutenberg", data-rich world.
   • There has been an efflorescence of meetings and discussions amongst disciplinary and informatics groups on the issue.
   • Some learned societies are well advanced (e.g. the RSC), but others less so, in taking up the issue.

3. On the negative side, HEFCE is, at the moment, not inclined to include data in the REF after next. The EU data protection legislation may place serious and unnecessary constraints on open data for science, and inhibit a liberal TDM regime. However, a copyright reform bill designed to except the UK from the latter will be presented to Parliament in its next session.

4. Other international moves are very positive, including the G8 open science statement, which included open scientific research data, the creation of the Research Data Alliance (US-EU-Australia), changes at the top in CODATA, the European Commission's Open Data Strategy and an Horizon2020 open data pilot.

A Forum

5. In the UK, there is a great deal of energy and creativity at grass roots level, and buy-in to the concept of the power of open data at a high level. The purpose of a Forum would be to explore how those with levers for action in their hands (universities, RCs, FCs, publishers) can best collaborate to overcome the barriers and create enabling mechanisms that will facilitate a powerful open data regime, and to ensure that the UK is a leader in a new wave of open science.

6. The agenda should include an analysis of how the key institutional recommendations in the report could be addressed and how barriers to implementation could be removed. These were:
Universities & Institutes: open data as default; data communication as a criterion for progression & reward; developing a data strategy.
Funding Councils: assess open data deposition on the same basis as publication; reward collaborative working.
Learned Societies (in their communities, to promote): open data; enhance data management; advocate open data principles.
Research Councils & Charities: mandate "intelligent openness in projects they fund; recognize those who maximize data usability and communication; data deposition and metadata preparation as a cost of research.
Scientific publishers: mandate intelligent openness.

7. Other major issues for the Forum:
   - Creating and characterising a consensus about importance of the issue.
   - Creating a data taxonomy (for example, learned societies, which tend to express the priorities and standards of their disciplines would be able to identify the categories of data with which they deal, which when collated across the disciplines would provide a vital view of the data landscape and the opportunities/problems it offers).
   - Agreeing standards for metadata and exploring standards of collection at a subdisciplinary level.
   - The relationship between individual and collective action, and the role of other databases (community – e.g. chemistry – national/international) in data curation in universities/institutes.
   - Identifying the need for technical informatics solutions, signaling these priorities to funders of research, and the role of international collaboration.
   - Ensuring that the advice from data scientists is fed into discussions.
   - Exploring the boundaries – (commercial interests, confidentiality, safety and security)
   - Identifying possible roles for Government.
   - Identifying the need for skills.
   - Identifying priorities for a national data infrastructure and its relations to other European and international projects.

A Forum Meeting

8. The meeting should consist of representatives of groups identified in the report and that are able to contribute to its objectives. It should be a 1.5 day meeting. It will be held ayt the Royal Society. Representatives of the following should be invited:
   - Universities (UUK/Russell Group)
   - RCUK
   - Universities Funding Council
   - Learned societies

- ➢ Scientific publishers
- ➢ British Library
- ➢ Business spokespersons

Programme

Kick-off
Welcome
　　Talk 1: Setting the Open Science agenda
　　Talk 2: The commercial interface
　　Talk 3: Open data in practice (Royal Society of Chemistry, European Bioinformatics Institute)

Talks to be inserted in the agenda at appropriate times:
　　Talk 4: Personal privacy
　　Talk 5: Safety & security
　　Talk 6: Informatics solutions

Discussion menu: Main menu of discussions to focus on the issues in 6 & 7 above, each led by a member of the Open Data working group.