The BBSRC and Data Sharing

In May 2007, BBSRC formally launched its new Data Sharing Policy (DSP). The policy states that research data generated as a result of BBSRC support should be made available with as few restrictions as possible, in a timely and responsible manner, to the scientific community for examination and use. The policy aims to achieve the sharing of data in both a timely and scientifically appropriate manner.

The policy is positioned at a high level and is potentially applicable to all research areas supported by BBSRC. Importantly, it does not prescribe one particular approach to data sharing but recognises that different approaches are appropriate to different fields of research. In taking this position, it accommodates pre-existing data sharing activities (e.g. in the biomolecular sciences and sequencing) and provides guidance on good practice for emerging areas of research (e.g. systems biology).

Key principles of the DSP statement are:

- Data are important not only for the researchers generating the data but also to the wider scientific community, which may wish to examine and use these datasets to underpin other investigations;
- Data sharing should be driven by scientific need and cost effectiveness;
- Timely release should generally be no later than the release through publication of the main findings and should be in-line with established best practice in the field. Where this doesn't exist, release within 3 years of dataset generation is suggested as a guide;
- The need to safeguard Intellectual Property and to protect opportunities for commercialisation of research outputs together with recognition that this should not unduly delay or prevent data sharing;
- Recognition that currently there are two areas where there is a particularly strong case for sharing data (a) data arising from high volume experimentation; and (b) low throughput data arising from long time series or cumulative approaches;
- Existing standards and established community resources should be used where possible to share data for example *via* existing databases or repositories, and researchers are encouraged to share data through mechanisms affording the widest availability for generating added value and enabling re-use.

Key to implementing the policy, all applicants to BBSRC must now include data sharing and management plans as part of research grant proposals and to request the associated funds under full economic costing. At present the policy does not extend to studentships and fellowships, although implementation via these routes is under consideration.

Data sharing statements are considered alongside the scientific case for support by referees and BBSRC assessment Committees or Panels. It is unlikely that a research grant proposal would fail on the basis of its data sharing policy. However, if a Research Committee considered that an otherwise excellent proposal had an inappropriate data sharing plan it could make a conditional award and / or provide feedback.

The published policy statement[1] is accompanied by comprehensive implementation guidance. This covers, in more detail, the key principles in the policy statement, the implementation route and the BBSRC funding mechanisms that exist to encourage and facilitate data sharing initiatives in the biosciences.

Following the formal launch of the DSP, BBSRC has established DSP Monitoring Group, Chaired by Professor Chris Rawlings (Rothamsted Research). This Group will be convened for a 24 month period initially and will (a) evaluate the policy's implementation and (b) determine whether the policy instrument is functioning effectively and appropriately.

Currently, BBSRC has two funding schemes that provide mechanisms to fund data sharing activities. These are listed below, including details of relevant awards made to date:

1. Bioinformatics and Biological Resources Fund

A pilot Bioinformatics and Biological Resources (BBR) Fund was launched in 2006 in pilot form, as part of BBSRC's Tools and Resources Programme. The BBR fund aims to support high quality, strategically relevant BBRs that are necessary to underpin the UK's international quality bioscience. Arising from the pilot call, £6.4M has been invested to fund 10 community resources. Six of these funded projects have now started:

- Web Services 4 Life Sciences (WS4LS): WS4LS aims to provide easy access to a catalogue of Web Services relevant to the life sciences communities. (PIs: Carole Goble, Manchester & Rodrigo Lopez, EMBL - European Bioinformatics Institute).
- Wheat Functional Genomics Resource: This resource will organise the UK's wheat tools and resources such that they become accessible to all. (PI: Keith Edwards, Bristol).
- RevGenUk: Reverse Genetics in Dicots: Integrated reverse genetics platform for the major UK dicotyledonous models will include mutagenised populations of legumes and *Brassicas* collated in an open source database with an integrated user-friendly web browser. (PI: Trevor Wang, John Innes Centre).
- FlyAtlas: Comprehensive and authorative Affymetrix micro-array based atlas of gene expression in different life stages of *Drosophila*. (PI: Julian Dow, Glasgow)
- Protein Circular Dichroism Data Bank (PCDDB): Internationally unique database to create, curate, and maintain a comprehensive electronic archiving and analysis resource for CD spectroscopy for scientific community. (PIs: Bonnie Wallace, Birkbeck College & Robert Janes, Queen Mary, University of London)
- BioModels: will provide the UK Researchers and the world a reference resource for storing, annotating and sharing quantitative models of biological interest. (PI: Nicolas Le Novere, EMBL - European Bioinformatics Institute).

A new BBR call was launched in June 2008, with a closing date for Expressions of Interest on 31 July[2]. A new series of awards is anticipated to be made in Spring 2009.

[1] http://www.bbsrc.ac.uk/publications/policy/data_sharing_policy.pdf
[2] http://www.bbsrc.ac.uk/funding/opportunities/2008/bioinformatics_biological_resources.html

2. Tools and Resources Development Fund

The Tools and Resources Development Fund (TRDF) was launched late 2005, as part of BBSRC's Tools and Resources Programme. The TRDF is intended to support small or short-duration, pump priming research projects and / or to bring together communities for collaborative purposes. One of the highlighted areas for support is for projects to facilitate data sharing within the bioscience community (for example in genomics, proteomics, metabolomics and systems biology) including the development of data standards and / or software. Examples of broadly relevant awards made in this area to date (non-exhausive) include:

- Omics Data Standards: synergy and implementations. (PI: Suzanna Sansoni, EMBL-EBI);
- Software infrastructure to support the standard of model curation and annotation (MIRIAM). (PI: Nicolas Novere, EMBL-EBI);
- Development of integrated web interfaces for Bioconductor genomic data analysis, annotation and visualization tools. (PI: Alvis Brazma, EMBL-EBI);
- Development of network analysis tool BioLayout Express3D. (PI: Thomas Freeman, Univ. Edinburgh);
- Informatics tools for analysis of quantitative proteomics data. (PI: Simon Hubbard, Univ. Manchester);
- THE FLYDATA PROJECT: Decision Support and Semantic Organization of Laboratory Data in Drosophila Gene Expression Experiments. (PI: David Shotton, Univ. Oxford).