Biotechnology and Biological Sciences Research Council (BBSRC): BBSRC Template

Data areas and data types

Outline the volume, type and content of data that will be generated e.g. experimental measurements, models, records and images

Guidance:

BBSRC recognises that effective data sharing is already practiced in certain areas and expects this to continue. BBSRC supports, either directly or indirectly, a number of such resources. Data sharing in other areas is also expected where there is a strong scientific case and where it is cost effective.

BBSRC has identified a number of areas where there is a particularly strong scientific case for data sharing. These are:

- Data arising from high volume experimentation
- Low throughput data arising from long time series or cumulative approaches
- Models generated using systems approaches

BBSRC expects data sharing to take place in these areas.

Standards and metadata

Outline the standards and methodologies that will be adopted for data collection and management, and why these have been selected

Guidance:

Standards are fundamental to effective data sharing. These can include standards for administrative processes, as well as for methodologies relating to data management and data formats. Researchers are expected to make use of current guidance and information on best practice.

It is expected that, in order to maximise the potential for re-use of data, BBSRC researchers should generate and manage data using existing widely accepted formats and methodologies where available. Data released for sharing should be validated and verified in line with accepted best practice and be of high quality. Data should be accompanied by the contextual information or documentation (metadata) needed to provide a secondary user with any necessary details on the origin or manipulation of the data in order to prevent any misuse, misinterpretation or confusion. Where standards for metadata exist, it is expected that these should be adhered to.

BBSRC encourages community development of standards where these do not currently exist or are not widely accepted and provides funding mechanisms for support of this type of activity.

Relationship to other data

State the relationship to other data available in public repositories

Secondary Use

Outline the further intended and/or foreseeable research uses for the completed dataset(s)

Guidance:

BBSRC supports the view that those enabling sharing should receive full and appropriate recognition by funders, their academic institutions and new users for promoting secondary

research.

Where data are shared through a third party resource or databases, secondary users should acknowledge the source of data. Where data are shared directly from the originator, depending on the level of usage and collaboration either joint authorship or acknowledgement to the data originator may be appropriate. It is also important to ensure that researchers and their research institutions are protected against claims that application of their data led to wrong conclusions/decisions by others: any use made of any data generated by third parties would not come with a warranty of its quality.

Furthermore, BBSRC expects that researchers accessing data have responsibilities to preserve data confidentiality and to observe the ethical and legal obligations pertaining to the data.

Methods for data sharing

Outline the planned mechanisms for making these data available, e.g. through deposition in existing public databases or on request, including access mechanisms where appropriate

Guidance:

BBSRC recognises that different approaches to data sharing will be required in different situations and considers that it is most appropriate for researchers to determine their own strategies for data sharing and outline these within their research grant proposal(s). Applicants should consider where, how, and to whom their data should be made available.

In addition, data sharing practices will change as areas of research develop and become more mature. This can be observed by looking at the areas of sequencing (i.e. well established mechanisms in place), microarrays (i.e. standards developed and being implemented) and systems biology (i.e. databases currently not well developed). Consideration should be given to what constitutes good practice in emerging areas of research.

It is expected that data sharing strategies will fall into the two broad categories below.

Data Sharing via a 3rd Party

Data sharing via deposition in an existing database, repository or other community resource is expected where possible and researchers are encouraged to share data through mechanisms affording the widest availability for generating added value and enabling re-use.

Researchers are encouraged to use existing infrastructure to facilitate data sharing where possible. BBSRC funds or otherwise supports a number of such resources. Where no such resources exist, applicants may consider sharing data via other third party mechanisms such as journal websites and / or open access repositories, many of which are now able to capture and share data underpinning publications.

Direct Data Sharing: from Originator to Others

This method of data sharing may be appropriate for areas where suitable third party mechanisms are not available. Researchers are expected to ensure that data are maintained for a period of 10 years after the completion of the research project in suitable accessible formats using established standards where possible such that the data can be made available on request in line with BBSRC guidance on good scientific practice. This may lead to collaboration between the new user and the original data creators, with the responsibilities and rights of all parties agreed at the outset.

Other mechanisms for data sharing may be used where appropriate. These could include sharing data within closed communities or a combination of methods for different datasets. Specific access mechanisms could be appropriate for example where there are ethical considerations, a need to protect confidential data, or other reasons for limiting access.

Proprietary data

Outline any restrictions on data sharing due to the need to protect proprietary or patentable data

Guidance:

In instances where BBSRC and a commercial partner jointly fund academic research work (for example LINK projects) there may be some restrictions over releasing data. Any such restrictions on data sharing due to co-funding arrangements should be set out in the “statement on data sharing” section of an application and will be considered when a grant application is peer reviewed. Applicants should also ensure they have obtained necessary clearances from relevant collaborators with regards to the content of the proposal including the data sharing plan in line with the BBSRC Research Grants Guide.

Timeframes

State the timescales for public release of data

Guidance:

The value of data often depends on timeliness. Researchers have a legitimate interest in benefiting from their own time and effort in producing data, but not in prolonged exclusive use of these data. BBSRC expects that all data (with accompanying metadata) should be shared in a timely fashion as soon as it is verified. It is expected that timely release would 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 best practices does not exist release within three years of generation of the dataset is suggested as a guide.

The timescale for release for the data may differ for several reasons, depending on the nature of the data. These reasons may include:

- **Scientific Area:** Researchers are expected to make data available in-line with established practices within the relevant research community. Examples include:
    - Crystallography (Protein Data Bank) - the community has agreed a maximum 12-month delay between publishing the first paper on a structure and making coordinates public for secondary use.
    - Sequencing (EMBL Nucleotide Sequence database) – submitted data can be withheld from public access until publication of results but no later.
    - Metabolomics (MeT-RO) – Up to a six-month delay in publication can be requested.
    - Arabidopsis microarray data (NASC Affymetrix service) – all data are made available after a maximum one-year confidential period.
- **Intellectual Property (IP) issues and potential for commercialisation of research outputs:** New knowledge generates patentable ideas. BBSRC is also driving a policy of Knowledge Transfer and strongly encourages the commercialisation of IP through various initiatives. BBSRC recognises the need for periods of exclusive use of data but considers that commercialisation of research does not preclude data sharing and should not unduly delay or prevent data sharing. Any IP issues or plans for commercialisation should be highlighted in the case for support of the grant application.
- **Length or scope of research project:** Data from large studies may be released in waves as they become available or as they are published.

Formats

State the format of the final dataset