FASEB
Federation of American Societies for Experimental Biology

Representing Over 125,000 Researchers

301.634.7000
www.faseb.org

9650 Rockville Pike
Bethesda, MD 20814

March 1, 2016

FASEB Statement on Data Management and Access

The Federation of American Societies for Experimental Biology (FASEB) affirms the importance of data management and access to scientific progress. Good data practices are necessary to achieve the maximal benefit of research for all stakeholders. Technological advances are expanding the ways investigators collect, utilize, and share data, leading to new knowledge and discovery. The diversity of data types, research areas, and resources available make it challenging to identify data management and accessibility strategies that are practical and relevant for all life science fields. Therefore, FASEB advocates for flexible and customizable approaches that allow investigators and research sponsors to establish reasonable expectations for a particular research project. Moving forward, the scientific enterprise will need to develop integrated community-based solutions. The following principles and recommendations are meant to help guide stakeholder efforts to advance data management and access in the biological and medical sciences.

Guiding Principles

- Improving data management and increasing data access can create new scientific opportunities
- Efforts to increase data access should consider the infrastructure required for data management, standardization, discovery, access, citation, reuse, sustainability, and long-term preservation – all of which are necessary for productive data sharing
- Expansion of requirements for data management and access will require commensurate financial and staff support from research sponsors
- Flexibility and adaptability are essential for any data management or access policies; the varied and rapidly evolving data landscape necessitates customized strategies
- Efficient and enabling data access requires a balanced approach that prioritizes datasets of high potential utility; policies should recognize that access to some datasets may not be worth the cost of sharing and long-term preservation
- Regulatory and administrative burden should be minimized so that any requirements ultimately promote science rather than hinder research

I. Data Management Plans

Data management plans (DMPs) are an important tool for promoting quality data management and appropriate data access. They can serve as a helpful planning exercise at the beginning of a project and focus attention on data sharing goals. Submission of a DMP can clarify expectations between investigators and their research sponsor. Flexibility and adaptability can be achieved by having individual investigators develop a DMP specific to their research area, data types used, and resources available.

Research sponsors may also enlist DMPs for secondary uses of benefit to the research community, such as identifying common resource needs and other barriers.

1. DMP Requirements: To attain the benefits of DMPs without creating unnecessary burden, DMPs should be short summary documents that address the most essential aspects of data management and access. In most cases, one to two pages should be sufficient, although additional information could be requested *just-in-time* for select circumstances. FASEB recommends the following DMP content requirements:

a. Description of the data and metadata to be collected
b. Overview of data management practices
c. Summary of any data sharing restrictions (confidentiality, intellectual property, etc.)
d. For **shared data**, information about when it will be made available, where it will be stored, how it will be maintained, and how others will be able to find, access, and reuse it
e. For **data that will not be shared**, justification for not making it accessible (which may include considerations of feasibility, data utility, etc. as well as sharing restrictions)

2. DMP Compliance Reporting: It will take time for funding sponsors and research communities to establish what constitutes reasonable practices and expectations for the many different areas of research and types of data. Therefore, FASEB recommends that research sponsors delay any DMP enforcement actions for the first five years. During this time, sponsors should:

a. Identify and address common barriers and emerging problems
b. Establish a process for modifying or updating DMPs
c. Ensure there is sufficient flexibility and adaptability built into all requirements
d. Standardize policies and reporting requirements with other sponsors (particularly among federal funding agencies)
e. Outline the roles and responsibilities of all parties for data management after the grant ends

Once sponsors establish harmonized and well-vetted DMP policies, continuous assessment will be necessary to ensure that the policies do not delay the adoption of improved practices or new technologies.

II. Roles of Stakeholders in Improving Data Management and Access

FASEB recognizes that improving data management and access is an important and evolving challenge for the research community. As science and technology advances, so must data practices. Research sponsors, investigators, institutions, and scientific journals can contribute to and benefit from advancing data management and access strategies.

A. Research Sponsors

1. Incentives: Sponsors should encourage investigators to improve management of and access to their research data. This should include actions that maximize the value of shared data and assure professional recognition for making datasets accessible. As a first step, research sponsors should provide investigators with resources and services, including:

a. Creation of and continued support for additional databases and repositories
b. Help identifying any relevant databases, repositories, and other resources
c. Development of training modules on data management and sharing
d. Assistance in procuring a unique digital object identifier (DOI) for shared datasets

In the long-term, research sponsors should also address the resources and measures needed to promote productive data sharing:

a. Development of a single unified portal system to discover datasets and a unified system for metadata submission to data catalogues
b. Advancement of data citation practices, including the provision of standardized, exportable citation information for datasets included in the sponsor's data catalogues or databases
c. Facilitation of the development of community-based data standards by convening stakeholders, as needed. Scientific societies can assist by identifying experts, providing thoughtful feedback, and disseminating proposed standards
d. Provision of long-term data storage options when no relevant database exists (such as sponsor-based databases or "dark" storage to serve as a back-up)

2. Policies and Rules: Research sponsors also can use policies and rules to ensure that data produced through supported research are appropriately managed and made accessible. FASEB affirms that a sponsor's expectations should be commensurate with the resources available to the investigator and the sponsor's own support for such resources. At a minimum, sponsors should provide sufficient support to fully comply with all applicable data management and access requirements *as part of a project's funding*.

To effect positive change, research sponsors must also carefully balance the costs and benefits of data access when developing and amending policies. Making datasets accessible – including the skilled human labor necessary to prepare and maintain data and metadata, technological infrastructure, and continued development of effective search platforms – is costly. Some datasets have little value for reuse or a short "shelf-life"; requirements to share and preserve such data could create inefficiencies in research funding and resource distribution. FASEB recommends that sponsors ensure their data access policies prioritize data with the highest potential for reuse.

B. Investigators

3. Data Management: Quality data management is an essential component of productive data sharing. Poor practices can render a potentially valuable dataset useless. At a minimum, FASEB recommends that investigators ensure the following data management practices are established and maintained within their own laboratory:

a. Regular back-up of ***digital*** data onto a well-maintained server, cloud, or separate machine (ideally an automated process utilizing offsite backup storage)
b. Standardized meta-data collection and documentation for ***common*** data types used or produced within the laboratory
c. Sufficient documentation to facilitate dataset retrieval several years after collection
d. When possible, use of unique identifiers in metadata fields (e.g., an ORCID iD for individuals)
e. Prompt training of all research team members on the laboratory's data management practices

4. Data Sharing: If there are no restrictions or other considerations that would preclude sharing, investigators should submit ***key*** data from their research to a relevant database or repository. If no publicly-accessible topical or data type-based repository exists, investigators should establish plans for making the data available. This might include sharing upon request, publishing supporting data in the supplemental materials of an article, or depositing data into a non-specific database.

C. Scientific Journals

5. Role in Compliance: The point of publication occurs too late in the research process to effectively address many issues related to good data practices. Furthermore, most journals do not have the capacity to confirm author compliance with any applicable DMPs and policies. Therefore, FASEB strongly recommends that the federal government and other research sponsors avoid requiring journals to assure compliance with DMPs, confirm data are and remain accessible, or provide database services. Such activities would be more effectively and efficiently managed by the sponsor and grantees.

6. Professional Norms: FASEB affirms the role of scientific journals in promoting good practices and encourages them to: (1) request that authors include the DOIs for and/or web addresses of datasets in their original manuscript; and (2) uphold that, in the absence of any restrictions on sharing or similar concerns, investigators are responsible for making the underlying data available upon request.

D. Research Institutions

7. Resources: Institutions should provide the technological infrastructure necessary for quality data management and compliance with DMPs. At a minimum, investigators must be able to attain the professional data norms within their field of research.

8. Professional Culture: Institutions should also foster an atmosphere where of quality data management and appropriate data sharing are standard practice. To establish and maintain such an environment, FASEB recommends that institutions ensure the following:

a. Appropriate data training is available for all individuals conducting research
b. Institutional resources for data management can be easily identified and utilized
c. Investigators are encouraged to collaborate on improving data practices at the institution and within their discipline.