**Strategies for Financial Sustainability of Publicly Funded Research Databases**

***Asking the Right Questions***

**Board on Research Data and Information**

**National Academy of Sciences**

Over the past couple of decades, the research community has increasingly agreed that data are both a valuable product as well as a key resource for scientific progress. Data-driven discovery has become a new research paradigm. The growing production of data, the more extensive applications and reuse of those data, and the recognition of the value of data as a research resource have all led to a realization that new infrastructure and management approaches need to be developed to preserve and further exploit such data. Moreover, new and existing policy and guidelines for the inclusion of data management plans within proposals is an important recognition at the federal level that digital data are a fundamental component of 21st century research. Discussions on the data management plan policy are increasing the engagement between federal agencies and the research community. These discussions can shed light on the needs for access and preservation of valuable digital data as part of the research enterprise, as well as the need to plan throughout the digital data life cycle to accommodate the use, re-use, and possible stewardship of research data.

While the nation faces a critical need for scientific data preservation, repositories, and reuse and repurposing of many of the data already collected, we also confront various challenges, particularly economic and socio-cultural ones, to make better use of these public investments. Improvements in scientific data sustainability practices and policies could contribute very significantly to the economic and social progress of the nation in the context of the “knowledge economy” and “information society.” The White House Office of Science and Technology Policy memorandum of February 22, 2013 on Increasing Access to the Results of Federally Funded Scientific Research expressly recognizes these issues.

In order to best meet these needs, it makes sense to examine both the supply and the demand sides of the problem. Although not all data can or should be preserved for the long-term, strategies for appropriate long-term access and sustainable preservation, as well as the use and re-use of research data are important not only for promoting innovation, discovery, and the progress of science more generally, but are essential for capturing the return on public investments in research more fully.

The purpose of this one-day workshop, therefore, is to sharpen the questions that need to be addressed in this context and to structure a study that would be conducted in the subsequent months by the Board on Research Data and Information. The workshop would include sessions of invited speakers and an open and invited expert audience in an interactive format to explore the issues of: (1) Scientific Data as Research Infrastructure (Types of repositories/taxonomy of data types and institutions in different sectors/user requirements and principles); (2) Business Models and Economics of Sustainable Data Infrastructures (Examples/issues/principles); (3) Session Three—Current Policies for Data Sustainability (with a focus on existing and planned policy framework); and (4) Session Four—Asking the Right Questions (What needs to be known to make progress in this area and implement a long-term data infrastructure?). The presentation slides and audio transcript would be posted openly and a written transcript would be made available to the sponsor.