Sustainability Strategies for Publicly Funded Scientific Databases
Consensus Study

Board on Research Data and Information
Policy and Global Affairs Division
National Academy of Sciences

Draft Proposal

SUMMARY
The National Research Council's (NRC's) Board on Research Data and Information (BRDI) proposes to establish an ad hoc study committee to conduct a study whose goal is to characterize and provide sustainability strategies for valued scientific data resources produced primarily through federally funded research. The charge to the study committee will include the following tasks:

1. Assess the research data landscape by identifying and characterizing the types of publicly-funded scientific data of broad interest and use by research, education, commercial and professional communities today, and also explore how the data resources are likely to change over the next decade, by

   a. Developing a core set of common-use case models that characterize the size, location, longevity, need for access and preservation, required services and tools, and other characteristics considered important for publicly-funded scientific data; and

   b. Within each of these types of data and areas, providing exemplars of the value propositions and selection criteria that might be used to determine digital data of long-term value.

2. Examine the existing approaches and costs for long-term preservation and access for federally-funded digital scientific data (described in Task 1, above), and provide a gap analysis between existing options for data stewardship and community needs.

3. Provide conclusions and recommendations to the sponsors regarding an overall sustainability strategy, projecting 10-years out, including:

   a. Suggestions for federal programs that can be used to close the gap between projected levels of valuable digital data (as described in Task 1) and available options for data stewardship (as analyzed in Task 2)

   b. Potential vehicles for public, private, and academic solutions, as well as partnerships among these sectors, that could address the gap between the need for stewardship of valuable digital scientific data and sustainable stewardship options, and

   c. Potential approaches to integrate key requirements of the digital data life cycle into funding instruments, infrastructure programs, and other suitable mechanisms.

The study will be performed in 22 months and the resulting consensus study will be published in accordance with NRC procedures.