Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age

Board on Research Data and Information

January 30, 2009

Tom Arrison, Study Director

Context for the Study

- A sea-change in digital data and large data collections in science and engineering
- Policy making is increasingly data-driven and complex (e.g. climate change, environment, drug approval)
- Concerns about integrity (stem cell scandal, digital image manipulation)
- Differences in “data cultures” between fields (e.g. expectations regarding openness and sharing, etc.)

Origin

- Journals and others ask that the Academies take up data issues
- Committee organized under COSEPUP spring 2007
- Sponsored by the Academies, journals, societies, federal agencies, private companies

Committee

Daniel Kleppner, Lester Wolf Professor of Physics Emeritus, MIT (Co-Chair)
Phillip A. Sharp, Institute Professor, Center for Cancer Research, MIT (Co-Chair)
Margaret A. Berger, Professor of Law, Brooklyn Law School
Norman M. Bradburn, Professor Emeritus, University of Chicago
John Brauman, J.G. Jackson - C.J. Wood Professor, Stanford University
Jennifer T. Chayes, Principal Researcher, Microsoft Research
Anita Jones, Professor of Engineering and Applied, Science, University of Virginia
Linda P.B. Katehi, Provost, University of Illinois, Urbana-Champaign
Neal F. Lane, University Professor and Senior Fellow, Rice University
W. Carl Lineberger, Professor of Chemistry, University of Colorado
Richard Luce, Vice Provost and Director of University Libraries. Emory University
Thomas O. McGarity, Chair, Trial & Appellate Advocacy, University of Texas, Austin
Steven M. Paul, Executive Vice President, Science & Technology, Eli Lilly and Company
Teresa A. Sullivan, Provost, University of Michigan
Michael S. Turner, Professor, Department of Astronomy and Astrophysics, University of Chicago
J. Anthony Tyson, Distinguished Professor of Physics, University of California, Davis
Steven C. Wofsy, Professor of Atmospheric and Environmental Sciences, Harvard University

Questions to be addressed - I

1. What are the growing varieties of research data? In addition to issues concerned with the direct products of research, what issues are involved in the treatment of raw data, pre-publication data, materials, algorithms, and computer codes?

2. Who owns research data, particularly that which results from federally-funded research? Is it the public? The research institution? The lab? The researcher?

Questions to be addressed - II

3. To what extent is a scientist responsible for supplying research data to other scientists (including those who seek to reproduce the research) and to other parties who request them? Is a scientist responsible for supplying data, algorithms and computer codes to other scientists who request them?

4. What challenges does the science and technology community face arising from actions that would compromise the integrity of research data? What steps should be taken by the science and technology community, research institutions, journal publishers, and funders of research in response to these challenges?

Questions to be addressed - III

5. What are the current standards for accessing and maintaining research data, and, how should these evolve in the future? How might such standards differ for federally-funded and privately-funded research, and for research conducted in academia, government, nongovernmental organizations, and industry?

Study Timeline

- Committee meetings in April 2007, September 2007, December 2007
- Draft report April 2008, goes through several rounds of revisions, sent to reviewers in December
- Almost all reviews in, report release expected March 2009

Key Points

- Principles for data integrity, accessibility, and stewardship
- Suggest roles and recommendations for researchers, research institutions, sponsors, journals, and professional societies
- Focus on enabling greater openness and transparency

Possible Outcomes and Next Steps

- Continued dialogue and cooperation among research enterprise components
- Encourage/help research communities to transition
- Specific steps to implement principles, define roles and responsibilities
- Regional workshops, education and training needs, etc.